

```
Db      668  CCCACAGACTATGGAACAGACGCTCTTATGTACACAGCGCGGAGGGTTGCGAC 727
Qy      121  GlyThrTyrGlnThrAsnValValArgSerGlyGlySerCysLeuTyrValProPro 140
Db      728  AGCAGCTATCCACGAAAGCGGTGTGGGAAACAAACGCGCTGCTGTGTACGTCGCCGCC 787
Qy      141  GlyIlePheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHis 160
Db      788  GGCATCTTCAAGACCACTGCAGATCGACATCACCTGGTTCCTCCCTTCGACGACCAACGA 847
Qy      161  CysAspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLys 180
Db      848  TGCAGATGAAGTTTGGCAGCTGACCTATGATGGTTATCAGTTGGATCTACAACCTACAG 907
Qy      181  AspGluAlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGly 200
Db      908  GATGAAGGGGCGGAGATATAAGCAGTTTGTACGAAATGGCGAATGGGAGTTAATAGGA 967
Qy      201  MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGluProTyrValAspVal 220
Db      968  GTCCCCGGCAAGCCACAGATCTACTACAACCTGTTGTTCGGAGCCATCATCGACATC 1027
Qy      221  ThrPheThrIleMetIleArgArgThrLeuTyrTyrPhePheAsnLeuIleValPro 240
Db      1028  ACGTTTCGGTGGTGATCCGGAGGAAACGCTCTACTACTTCTTCAATCTGATCGTCCGC 1087
Qy      241  CysValleuIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 260
Db      1088  TGGGTGCTCATCGCCTCCATGGCTCTATTGGGGTTCCACTTGCCTCCAGACTCCGGAGAA 1147
Qy      261  LysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAla 280
Db      1148  AAGTTGCTTTAGGTGACGATATTAAGTCTGCTTGACGGTGTCTTCAACATGTTGGCGC 1207
Qy      281  GluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIle 300
Db      1208  GAGACGATGCGACGACGTCGGACGCGTGCCTTGTCTGGCAGCTACTTCAACTGCATC 1267
Qy      301  MetPheMetValAlaSerSerValValLeuThrValValValLeuAsnTyrHisArg 320
Db      1268  ATGTTTATGTTGGTCTCTCCGCTGCTCCACCATACTGATCTCAACTACCAACCCGCG 1327
Qy      321  ThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPheLeuGlnTrpLeuPro 340
Db      1328  CAGCAGACACTCAGAAATAGTATGATGATGCTGCTGCTGCTTTCCTTATGCTGCCG 1387
Qy      341  TrpIleLeuArgMetSerArgProGlyLysIleThrArgLysThrIleMetMetAsn 360
Db      1388  TGGGTGCTGCGCATGTCAACGCGCGGCTCGCGCAGACGCGCGCGCGCGGTA--- 1444
Qy      361  ThrArgMetArgGluLeuGluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnVal 380
Db      1445  CTTCCGCGCGCGGACCTGGAGCTGCGGAGCGCTCTCCAAAGTGGCTCTTACGCAAGTG 1504
Qy      381  LeuAspIleAspAspPheArgHisGlyProProProProAsnSerThrAlaSerThr 400
Db      1505  CTGCAGACTCATGACGACTCCGCCAC-----CCGACAGCGCAGCAG----- 1546
Qy      401  GlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgSerPheValArg 420
Db      1547  -----CCGCAATGCTGCCGATAC-----TACAGGGGGGT----- 1576
Qy      421  ProSerThrMetGluAspValGlyGlyLeuGlySerHis-----His 435
Db      1577  -----GAGGAGAAATGGCGGGGTGGCGCGCGCACAGTTGCTTTCGTTGTCGAC 1624
Qy      436  ArgGluLeuHisLeuIleLeuArgGluPheIleThrAlaArgMetLysLysAla 455
Db      1625  TACAGAGCTCTCCCTCATCTCGAAGAGATTAGATCATACAGATCAGATGCGCAGGAC 1684
Qy      456  AspGluGluAlaGluLeuIleSerAspTrpLysPheAlaAlaMetValValAspArgPhe 475
Db      1685  GACGAGATCGGACATTTCCGCGGACTTGGAGTTGCGCCGCACTGCTGTCGTCGACAGACTG 1744
```

```
Qy      476  CysLeuPheValPheThrLeuPheThrIleIleAlaThrValAlaValLeuLeuSerAla 495
Db      1745  TGCCTTATTATCTTACCTGTTTCACAATCATGCCACGCTAGCGTCTGCTGTCGCGC 1804
Qy      496  ProHisIleIleVal 500
Db      1805  CCACACATCATGGTG 1819
RESULT 12
AX009612          3700 bp      DNA      linear      PAT 06-SEP-2000
LOCUS             Sequence 3 from Patent EP0962528.
DEFINITION        AX009612
ACCESSION         AX009612
VERSION           AX009612.1 GI:9996844
KEYWORDS
SOURCE            Heliothis virescens (tobacco budworm)
ORGANISM          Heliothis virescens
                  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                  Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
                  Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE
AUTHORS           Adamczewski, M.D., Schulte, T.D. and Oellers, N.D.
TITLE             Nucleic acids encoding acetylcholin-receptor subunits from insects
JOURNAL           Patent: EP 0962528-A 3 08-DEC-1999;
                  BAYER AG (DE)
FEATURES           Location/Qualifiers
                   1..3700
                   /organism="Heliothis virescens"
                   /mol_type="unassigned DNA"
                   /db_xref="taxon:7102"
                   335..1825
                   /note="unnamed protein product"
                   /codon_start=1
                   /protein_id="CAC07500.1"
                   /db_xref="GI:9996845"
                   /db_xref="RENTREMBL:CAC07500"
                   /translation="MGRARRSLAAPAGLLLLCLLMPRGARCGVHEKRLHLLHLLDH
                   YNVLRFVNESDPLQSLFGLTMIQIIDVDEKNQLLITNWLKLEWMDNMLRWNTSDF
                   GGKDLRVPFPHLWKPDLVMSADGDFSTPTNVVRNNGSLVYPFGIFKSTCKI
                   DIWFPPDQRCMEKFGSTYDGYQLDLOLQDEGGDISSFVTNGEWELIGVPGKRNE
                   IYVCCPEYDITFAVVRRTLYFENLIVPCVLASWALLGFTLPPDSGKLSIG
                   VTLLSLTFLNVAETPATSDAVLLGTYFNCIMFVASSVSTILLINLYHRRHAD
                   THEMDSWIKCVFLWLPVLRMSRGSATPPPARPPPPDLELRBSKSLIANVLD
                   IDDFHPQAQQPCCRYRGGSEAGLAASFCFVDYELSLILKEIRVITDQMRKD
                   DEDADISRDWKFAMVVDRLCLIFLTITIIATLAVLLSAPHINVS"
ORIGIN
Alignment Scores:  7.27e-164      Length:      3700
Pred. No.:         1803.50        Matches:     347
Percent Similarity: 79.21%        Conservative: 53
Best Local Similarity: 68.71%     Mismatches:   78
Query Match:       68.31%         Indels:      27
DB:                6              Gaps:         7
US-09-303-232-6 (1-501) x AX009612 (1-3700)
Qy      2  AlaProMetLeuAlaAlaLeuAlaLeuLeuAlaLeuProValSerGluGln--Gly 20
Db      368  GCGCCCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 427
Qy      21  ProHisGluLysArgLeuLeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgPro 40
Db      428  TACCACGAGAGCGGCTACTGACCACTATTGGACCACTTACCACTACTGAGAGAGGCC 487
Qy      41  ValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnIleIle 60
Db      488  GTCGTCAACGAGAGACCGCGCTGCGAGCTCTCCTCGCGCTCAGCGTCAATGAGATCATC 547
Qy      61  AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsn 80
Db      548  GACGTGACGAGAGAACCAACGCTTTTATTAACAAACATCTGGCTAAACTAGAGTGAAT 607
```







```

/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="2L: 34E4-5"
/dev_stage="embryo"
1. .2907
/gene="nAcRalpha-34E"
/allele="A"
289. .2712
/gene="nAcRalpha-34E"
/notes="ion channel; neurotransmitter transmembrane
receptor"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha5
subunit"
/protein_id="AAMI3390.1"
/db_xref="GI:20152840"
/translation="MKNAQLKLTEDVDELWLAVRLAHCSSNFSSSSSTRTSSNORH
NQQLTTLQPSLSLKHHNLTASBOHNSQQQEPASKDEEDVANHGSRNDQDTHLOQLDSS
NMSPKTAATAAGATAGBATTCQPTNIRLCARKQRLRRRRKRPATNETIDIRKQQOL
SMPPFKRKSTDTYTPAATTSCTPATYMQCRASDNFSPISIRHDRVSTATFAWLH
VLQVLLSQOQHLVQORSVLLFRRIAATIAFISYLGFAAQLNSSSSSSSSSS
NNSQTQLNGINKHSWIFLLYLNLKAVCLAGHEKRLHLLDLPYNTLRRPVINES
DPQLSFLGLTMOIIVDEKNQLLVTVWLKLEWMDMLRWNTSDYGVKDLRIPRER
IKPDLVYNASADEGFDGTQYTNVVRNNGSCLYVPPGIPKSKTIDITWEPDPDQRC
EMKGSWTYDGFOLDLODETGDSDISSYVINGEWELLGVPGKRNIEIYNCCPEPYID
ITAIIRRLIYFFNLIIPCVLIASMALLGFTLPDPSGEKUSLGLVILLISLTVFIN
MVATPMATSDAVPLGLTYFNCIMFVASSVSTILNLNHYHRNADTHEMSWIRIVF
LCWPLTRMSRGRPLIEFPPTPCSDTSERKHQLSDVLEKRSKSLILANLVDI
DDDFRCHRCMTPOGTLPHNPAFVRYTVYGGDDGSGIGPIGSTMPDVAVHTTICKST
EYELGLTKIRFITDQLRKDECDNDIANDKFAAMVVDRLCILIFTFMFAILATIVL
LSAPHIIVS"
373
/variation
/gene="nAcRalpha-34E"
/notes="results in threonine to isoleucine substitution;
compared to B allele"
375
/variation
/gene="nAcRalpha-34E"
/notes="results in phenylalanine to isoleucine
substitution; compared to B allele"
/replace="C"
827
/variation
/gene="nAcRalpha-34E"
/notes="results in threonine to isoleucine substitution;
compared to B allele"
/replace="t"
1079. .1080
/variation
/gene="nAcRalpha-34E"
/notes="results in lysine to arginine substitution;
compared to B allele"
/replace="gg"
1081. .1086
/variation
/gene="nAcRalpha-34E"
/notes="results in deletion of asparagine and serine;
compared to B allele"
/replace="n"
1119.1120
/variation
/gene="nAcRalpha-34E"
/notes="results in serine, glycine, and asparagine
insertion; compared to B allele"
/replace="agcggaac"
1121
/variation
/gene="nAcRalpha-34E"
/notes="results in asparagine to serine substitution;
compared to B allele"
/replace="g"

1.02e-163 Length: 2907
Pred. No.: 1800.50 Matches: 348
Score: 77.90% Conservative: 61
Percent Similarity:

ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:

```

```

Best Local Similarity: 66.29% Mismatches: 69
Query Match: 68.20% Indels: 47
DB: 3 Gaps: 9
US-09-303-232-6 (1-501) x AF272778 (1-2907)
Qy 10 LeuLeuAlaLeuLeuProValSerGluGln-----GlyProHisGluLysArg 25
Db 1171 TTATTGATATATTGAATTTATCTGCTAAAGTTTTCCTAGCAGGATATCATGAAAAAGAGA 1230
Qy 26 LeuLeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSer 45
Db 1231 CTGTTACACGATCTTTTGGATCCCTTATATACACTAGACGTCCTGCTTCAATGAATCG 1290
Qy 46 GluProLeuGluValArgPheGlyLeuThrLeuGlnLeuLeuLeuLeuLeuLeuLeuLeu 65
Db 1291 GACCCGTTACAATTAAGCTTTTGGTTTAACTTTAATGCAAAATTAATCGATGTGGACGAGAAA 1350
Qy 66 AsnGlnLeuLeuLeuThrAsnLeuLeuTrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArg 85
Db 1351 AATCAATTTGCTAGTCACTAATGTGTGTTAACTTGAAGTGGACGACATGAATCTCCGC 1410
Qy 86 TrpAsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThrProAsnLysLeuTrp 105
Db 1411 TGGACACCTCCGACTATGCGGAGTTAAGATCTGGAATACCGCCGATCGCATCTGG 1470
Qy 106 LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThr 125
Db 1471 AAGCCGACCTGCTGATGTACAAAGTGGGATGAGGATTTGACGGCACCCTACACGACG 1530
Qy 126 AsnValValValArgSerGlySerCysLeuTyrValProProGlyIlePheLysSer 145
Db 1531 AAGTGTGTGTGGGAAACACGGCTCGTGTCTATACGTCGCGGGGATCTTCAAGTCG 1590
Qy 146 ThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMetLysPhe 165
Db 1591 ACGTGCAGATCGACATCAGTGGTTCCTTCGATGACGCGGTGCGAGATGAAGTTTC 1650
Qy 166 GlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyGly 185
Db 1651 GGCAGTTGGACCTACGACGGATTCACAGTGGATTACAAATTACAAGATGAATGCGCGGT 1710
Qy 186 AspLeuSerAspPheIleThrAsnGlyGluTyrTrpLeuLeuGlyMetProGlyLysLys 205
Db 1711 GATATCAGAGTACGTCGTCACGCGAGTGGGAACTACTGGGTGTGCGCGGCAACAGT 1770
Qy 206 AsnThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMet 225
Db 1771 AACGAGATCTATTACAACCTGCTCCCGGAACCCCTATATAGACATCACCCTTCGCATCATC 1830
Qy 226 IleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuLeuSer 245
Db 1831 ATCCGCGCAGCAACACTGTACTATTCTTCAACCTGATCATCCTTGTGTACTGATTGCC 1890
Qy 246 SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyLysLeuThrLeuGly 265
Db 1891 TCCATGCGCTTGTTCGGATTTCACCTGCGCCAGATTCGGGTGAAAATAATTATCGCTGGGT 1950
Qy 266 ValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGln 285
Db 1951 GTTACCATCTTGCTCTCGCTGACCGGTGTTCTGAATATGTTGCGGAGACAAATGCCGGCT 2010
Qy 286 ValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAla 305
Db 2011 ACATCCGATGCGGTGCCATTTGCTGGGTACATAATTCAAATTCGATATGTTTATGGTAGCT 2070
Qy 306 SerSerValValLeuThrValValValLeuAsnTyrHisHisArgThrAlaAspIleHis 325
Db 2071 TCATCCGTTGTGTCAACGATTTTAAATATTAAATATATCAATCATCGAAATGCTGATCGCAC 2130
Qy 326 GluMetProGlnTrpIleLysSerValPheLeuGlnTrpLeuProTrpIleLeuArgMet 345
Db 2131 GAAATGTCCGATGGATAGCATCGCTGTTTTTGTGCTGGCTCCATGGATATTGCGAATG 2190

```

```
Qy 346 SerArgProGlyLysLysIleThrArgLys-----ThrIleMetMetAsnThrArg 362
Db 2191 AGTGGCCAGGACGACCGCTGATCCTAGAGTTCCGACCGCCCTGTTCCGACACATCC 2250
Qy 363 -----MetArgGluLeuGluLeuLysGluArgSerSerLysSer 375
Db 2251 TCCGAGCGGAGACACCATGACTCTCCGACGTTGAGCTGAAAGAGCGCTCGTGGAAATCG 2310
Qy 376 LeuLeuAlaAsnValLeuAspIleAspPheArgHisGly-----ProProPro 393
Db 2311 CTGCTGGCCACGACTAGACATCGATGATGACTTCCGGCACAAATTGTGCCCCCATGACG 2370
Qy 394 ProAsnSerThrAlaSerThrGlyAsnLeuGlyProGlyCysSerIlePheArgThrAsp 413
Db 2371 CCCGGC-----GGACACTGCCACACACCCGGCTTTCTATCGCACGGTT 2415
Qy 414 PheArgArgSerPheValArgProSerThrMetGluAspValGly-----GlyGlyLeu 431
Db 2416 TATGGACAAGGC-----GACGATGGCAGCAATTGGGCCCAATT 2451
Qy 432 GlySer-----HisHisArg----- 436
Db 2452 GGCAGACCCGAATGCCGGATGCCGTACCCATCATAGTCATCAAAATCATCACTGAA 2511
Qy 437 ---GluLeuHisLeuIleLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAla 455
Db 2512 TATGAATTAGGTTTAATCTTAAGGAAATTCGCTTTATACTGATCAGCTACGTAAAGAT 2571
Qy 456 AspGluGluAlaGluLeuSerAspTrpLysPheAlaAlaMetValValAspArgPhe 475
Db 2572 GACGAGTGCAATGACATTCGCAATGATTGAAATTTGCAGCTATGGTCGTTGCAGACTG 2631
Qy 476 CysLeuPheValPheThrLeuPheThrIleAlaThrValAlaValLeuLeuSerAla 495
Db 2632 TGCCTTATCATATTACAATGTTCCGAATATTAGCCACATAGCTGTACTACTATCAGCA 2691
Qy 496 ProHisIleIleVal 500
Db 2692 CCACATATTATTGTC 2706
```

Search completed: May 9, 2004, 00:55:43  
Job time : 5552.12 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2004, 06:39:04 ; Search time 526.234 Seconds

(without alignments)  
4044.486 Million cell updates/sec

Title: US-09-303-232-6

Perfect score: 2640

Sequence: 1 MAPMLAALALLPVSEQ.....LFTIATVALLSAPHIIVQ 501

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-Q=/cgn2\_1/USPTO\_spool/US09303232/runat\_07052004\_101109\_23873/app\_query.fasta\_1.2261

-DB=N\_Geneseq\_29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blastm62 -TRANS=human40.cdi

-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09303232@cgn\_1\_1200 @runat\_07052004\_101109\_23873 -NCFU=6 -ICPU=3

-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_29Jan04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002s.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2640	100.0	3109	3 AAZ24477	Aaz24477 H. viresc
2	1968	74.5	1540	4 ABL07231	Abi07231 Drosophil
3	1803.5	68.3	3700	3 AAZ24476	Aaz24476 H. viresc
4	1609	60.9	2886	3 AAZ24475	Aaz24475 D. melano
5	1283	48.6	936	4 ABL13733	Abi13733 Drosophil
6	1258.5	47.7	1509	3 AAC58395	Aac58395 Human PRO
7	1258.5	47.7	1509	4 AAC90380	Aac90380 Wild-type
8	1258.5	47.7	1876	2 AAV12197	Aav12197 Human neu

9	1258.5	47.7	1876	2	AAZ48239	Neuronal
10	1258.5	47.7	1876	6	ABS54875	Human neu
11	1258.5	47.7	1876	6	ABV73248	Human neu
12	1258.5	47.7	1876	8	ADA10864	Human neu
13	1254.5	47.5	1590	2	AAV44687	V27AT var
14	1253	47.5	2769	2	AAT59196	Neuronal
15	1252.5	47.4	1509	4	AAC90385	Mutant hu
16	1248.5	47.3	1509	4	AAC90386	Mutant hu
17	1242.5	47.1	1509	4	AAC90387	Mutant hu
18	1229	46.6	1964	6	ABZ11298	Human pol
19	1226.5	46.5	2101	2	AAT59197	Neuronal
20	1080	40.9	1915	4	AAD20962	Caenorhab
21	1046	39.6	803	4	ABL07799	Drosophil
22	989.5	37.5	1908	2	AAT48236	Neuronal
23	989.5	37.5	1908	8	ADA10856	Human neu
24	982.5	37.2	1584	7	ADA83809	Human CHR
25	979	37.1	1416	6	AAC90382	Chimeric
26	946.5	35.9	1654	6	ABS54871	Human neu
27	946.5	35.9	1654	6	ABV73244	Human neu
28	943	35.7	2210	4	ABL10275	Drosophil
29	940.5	35.6	1757	2	AAV12200	Human neu
30	933	35.3	1743	2	AAT48232	Neuronal
31	933	35.3	1743	6	ABS54874	Human neu
32	933	35.3	1743	6	ABV73247	Human neu
33	933	35.3	1743	8	ADA10862	Human neu
34	933	35.3	2385	4	ABL11821	Drosophil
35	922.5	34.9	2277	2	AAV12199	Human neu
36	922.5	34.9	2277	2	AAQ90387	Alpha 2 s
37	922.5	34.9	2277	6	ABS54870	Human neu
38	922.5	34.9	2277	6	ABV73243	Human neu
39	922.5	34.9	2664	2	AAT48235	Neuronal
40	922.5	34.9	2664	6	ABK92165	Prostate
41	922.5	34.9	2664	8	ADA10854	Human neu
42	922.5	34.9	2664	9	ADC71170	Human 205
43	922.5	34.9	2666	9	ADC71168	Human 205
44	922	34.9	1503	4	AAD20961	Caenorhab
45	918.5	34.8	1896	6	AAL45867	Modified

#### ALIGNMENTS

RESULT 1

AAZ24477

ID AAZ24477 standard; cDNA to mRNA; 3109 BP.

XX AC AAZ24477;

XX AC AAZ24477;

DT 17-FEB-2000 (first entry)

DE H. virescens acetyl-choline receptor DNA from clone Hva7-2.

XX AC Acetyl-choline receptor; insect; insecticide; screening;

KW neurotransmission; plant protection agent; conductance; AChR; ds.

OS Heliothis virescens.

XX AC Acetyl-choline receptor; insect; insecticide; screening;

PN DE19819829-A1.

XX PD 11-NOV-1999.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PF 04-MAY-1998; 98DE-01019829.

XX PS Claim 1a; Page 19-22; 26pp; German.

XX CC This invention describes a novel nucleic acid (NA) encoding a nicotinic

CC acetyl-choline receptor (I) from insects which can be used as an

CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)

CC (also vectors containing it, its regulatory regions, and antibodies

CC directed against (I)-encoded proteins) are used to screen for: (a) plant

CC protection agents that alter conductance of AChR, potentially useful as

CC insecticides, or (b) genes which encode polypeptides that are involved in

CC formation of functionally related AChR in insects. (I) are also used to

CC isolate and characterize the specified regulatory regions and for

CC recombinant production of (II). This sequence encodes an acetyl-choline

CC receptor isolated from *Heliothis virescens*

XX SQ Sequence 3109 BP; 923 A; 610 C; 742 G; 834 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,56-272 Length: 3109

Score: 2640,00 Matches: 501

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 3 Gaps: 0

US-09-303-232-6 (1-501) x AA224477 (1-3109)

Qy 1 MetAlaProMetLeuAlaAlaLeuAlaLeuAlaLeuProValSerGluGlnGly 20

Db 95 ATGGCCCTATGTTGGCGGCTTGGCGTGTGCTTTGGTGGCGGTATCGGACCAAGT 154

Qy 21 ProHisGluLysArgLeuLeuAsnAlaLeuLeuAlaLeuTyrAsnThrLeuGluArgPro 40

Db 155 CCTCAGAGAGAGACTCTCTGAAACGGTGTCTGGGAACTACAAACACCTCGAGCGACG 214

Qy 41 ValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIle 60

Db 215 GTGGCCAAACGAGAGCGAACCGCTAGAGGTCAAGTTCGGCTTGGCTTGCAGCAAAATCATT 274

Qy 61 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleThrLeuSerLeuGluTrpAsn 80

Db 275 GACGTGACGAGAGAAATCAACTACTTATAACCAATATATGCTGTGCTGGAGTGGAAAT 334

Qy 81 AspTyrAsnLeuArgTyrPheAspSerGluTyrGlyValLysAspLeuArgIleThr 100

Db 335 GACTACACCTCAGGTGGAAACAGACGAGTATGGCGGGTCAAGGACCTCAGGATCAG 394

Qy 101 ProAsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 120

Db 395 CCCAACCAAGTTGTGAGCGGACGCTCTTATGTATATAATAGTGTGACGAGGGTTTGAC 454

Qy 121 GlyThrTyrGlnThrAsnValValArgSerGlyGlySerCysLeuTyrValProPro 140

Db 455 GGGACCTTACAGACCAACAGTGGTGTGAGACGGCGGCGAGTGTGCTGTAGTGCACCT 514

Qy 141 GlyIlePheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHis 160

Db 515 GGCATATTCAGAGACACATGCAAGATGGACATCGGTGGTTTCCCTTCGACGACCAACAC 574

Qy 161 CysAspMetLysPheGlySerThrThrTyrAspGlyAsnGlnLeuAspLeuValLeuLys 180

Db 575 TGTGATATGAAGTTCCGTAGTGGACATATGACGGCAATCAGTTGGATCTGGTCTAATAA 634

Qy 181 AspGluAlaGlyValAspLeuSerAspPheIleThrAsnGlyGluTrpTyrIleGly 200

Db 635 GATAGGCGGCGGCGATCTATCGGACTTCATAACAAATGGGAGTGGTATCTAATAGA 694

Qy 201 MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGluProTyrValAspVal 220

Db 695 ATGCCAGGCAAAAGAACACATAACATACCGGTGCTGCCCGGAGCCCTACGTGACGTC 754

Qy 221 ThrPheThrIleMetIleArgArgThrLeuTyrTyrPhePheAsnLeuIleValPro 240

Db 755 ACCTTCACCATCATGATAAGAACGACCTTGTACTACTTCTTCAACCTGATCGTCCG 814

Qy 241 CysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 260

Db 815 TGGGTGTGATCATCATGATGGCACTCTCGGCTTCACTGCGCAGACTCCGGAGAG 874

Qy 261 LysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAla 280

Db 875 AAATCACACTGGAGTCACTATTCTCTATCGTACGGTGTTCCTCAACTGGTAGCC 934

Qy 281 GluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIle 300

Db 935 GAGACCTTGCACAGAGTCTCCGACGTATCCCTCTGTAGGAGAGTACTTCAATTGTCATC 994

Qy 301 MetPheMetValAlaSerSerValValLeuThrValValLeuAsnTyrHisArg 320

Db 995 ATGTTTCATGGTAGCGTCTGTGTGTACTGACTGTGTGTGTACTCAATTACCACATCGA 1054

Qy 321 ThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPheLeuGlnTrpLeuPro 340

Db 1055 ACAGCTCATATACATGAATGCCACAGTGGATAAAATCAGTATTCTTACAAATGGTTGCCA 1114

Qy 341 TrpIleLeuArgMetSerArgProGlyLysLysIleThrArgLysThrIleMetMetAsn 360

Db 1115 TGGTACTGCAATGTGAGGCCAGGAGAGATCACCAAGAAAGACTATATATGATGAAC 1174

Qy 361 ThrArgMetArgGluLeuGluLysGluArgSerSerLysSerLeuLeuAlaAsnVal 380

Db 1175 ACGAGGATGAGGAGCTGGAACTGAAAGAGAGGTCGTGCAAGTCTTGTGCGCAATGTT 1234

Qy 381 LeuAspIleAspAspPheArgHisGlyProProProProProProProProProPro 400

Db 1235 CTAGATATTGATGATGACTTCAGACACCGCCCTCCGCTCTAAACAGTACTGCTCGACC 1294

Qy 401 GlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgArgSerPheValArg 420

Db 1295 GGGAAATTTGGACCTGGGTGCTCAATATTCGACGAGATTTCCGTGGTCTGTTCCGCT 1354

Qy 421 ProSerThrMetGluAspValGlyGlyLeuGlySerHisArgGluLeuHisLeu 440

Db 1355 CCGTCCAGATGGAAGAGCTGGCGGGCTGGGTAGCCACCATCGGAGCTGCACCTC 1414

Qy 441 IleLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAlaAspGluAlaGlu 460

Db 1415 ATACTGAGAGAGCTGCATTCATCACGCCAGGATGAAGAAGGCTGATGAGGAAGCCGAG 1474

Qy 461 LeuIleSerAspTrpLysPheAlaAlaMetValValAspArgPheCysLeuPheValPhe 480

Db 1475 CTGATCAGCGACTGGAAGTTTGTGCGATGGTGTGTATAGGTTTGGCTGTGCTGTTC 1534

Qy 481 ThrLeuPheThrIleIleAlaThrValAlaValLeuLeuSerAlaProHisIleVal 500

Db 1535 ACATTTTCACATCATTCGCGACAGTAGTGTCTGTATCGGACCGCATATCATCTGTG 1594

Qy 501 Gln 501

Db 1595 CAA 1597

RESULT 2

ABL07231

ID ABL07231 standard; cDNA; 1540 BP.

XX ABL07231;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 16175.

KW Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

OS Drosophila melanogaster.

XX







QY 401 GlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgSerPheValArg 420  
 Db 1547 -----CCGCAATGCTGCCGATAC-----TACAGGGGGGGT----- 1576  
 QY 421 ProSerThrMetGluAspValGlyGlyLeuGlySerHis-----His 435  
 Db 1577 -----GAGGAGAAATGGCGGGGTGGCGGCGCACAGATTGCTTCGGTGTGCAC 1624  
 QY 436 ArgGluLeuHisLeuLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAla 455  
 Db 1625 TACAGACTCTCCCTCAATCTGAGAGAGATTAGAGTCATCACAGATCAGATCGCGCAGGAC 1684  
 QY 456 AspGluGluAlaGluLeuLeuSerAspTrpLysPheAlaAlaMetValValAspArgPhe 475  
 Db 1685 GACGAGATCGGACATTCGCGGCACTGGGAAGTTGCGCGGCTAGCGTGTGGAGACTG 1744  
 QY 476 CysLeuPheValPheThrLeuPheThrIleIleAlaThrValAlaValLeuLeuSerAla 495  
 Db 1745 TGCCTATTATCTTACCTGTTTACATCATCGCCAGCTAGCGTGTGCTGTCCGCG 1804  
 QY 496 ProHisIleIleVal 500  
 Db 1805 CCACACATCATGGTG 1819

## RESULT 4

ID AAZ24475 standard; cDNA to mRNA; 2886 BP.  
 XX AAZ24475;

DT 17-FEB-2000 (first entry)  
 DE D. melanogaster acetyl-choline receptor DNA from clone Da7.  
 XX Acetyl-choline receptor; nicotinic; insect; insecticide; screening;  
 KW neurotransmission; plant protection agent; conductance; AChR; ds.  
 XX Drosophila melanogaster.

Key Location/Qualifiers  
 CDS 372..2684  
 FT /\*tag= a  
 FT /product= "acetyl choline receptor."

DE19819829-Al.

PD 11-NOV-1999.

XX 04-MAY-1998; 98DE-01019829.

XX 04-MAY-1998; 98DE-01019829.

XX (FARB ) BAYER AG.

PA Adamczewski M, Oellers N, Schulte T;

XX WPI; 2000-014207/02.

DR P-PSDB; AAY50814.

XX New nucleic acid encoding a nicotinic acetylcholine receptor from  
 insects, used to identify potential insecticides.

PS Claim 1a; Page 8-12; 26pp; German.

XX This invention describes a novel nucleic acid (NA) encoding a nicotinic  
 CC acetyl-choline receptor (I) from insects which can be used as an  
 CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)  
 CC (also vectors containing it, its regulatory regions, and antibodies  
 CC directed against (I)-encoded proteins) are used to screen for: (a) plant  
 CC protection agents that alter conductance of AChR, potentially useful as  
 CC insecticides, or (b) genes which encode polypeptides that are involved in  
 CC formation of functionally related AChR in insects. (I) are also used to

CC isolate and characterize the specified regulatory regions and for  
 CC recombinant production of (II). This sequence encodes an acetyl-choline  
 CC receptor isolated from *Drosophila melanogaster*

SQ Sequence 2886 BP; 942 A; 687 C; 640 G; 617 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 9.51e-162 Length: 2886  
 Score: 1609.00 Matches: 319  
 Percent Similarity: 71.62% Conservative: 57  
 Best Local Similarity: 60.76% Mismatches: 65  
 Query Match: 60.95% Indels: 84  
 DB: 3 Gaps: 10

US-09-303-232-6 (1-501) x AAZ24475 (1-2886)

QY 10 LeuLeuAlaLeuLeuProValSerGluGln-----GlyProHisGluLysArg 25  
 Db 1254 TTATTGATATATTGAATTATCTGCTAAAGTTGGCTAGCAGGATATCATGAAAGAGA 1313  
 QY 26 LeuLeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSer 45  
 Db 1314 CTGTTACAGCATCTTTTGGATCCTTATAATACACTAGAACGTCCTCAATGAATCG 1373  
 QY 46 GluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleIleAspValAspGluLys 65  
 Db 1374 GACCCGTTACAATTAAAGCTTTGGTTAACTTTAATGCAAAATTATCGATGTGGACGAAA 1433  
 QY 66 AsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArg 85  
 Db 1434 AATCAATTGCTAGTCACTAATGTGTGTTAAACTGGAGTGGACGACATGAATCTCCGC 1493  
 QY 86 TrpAsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThrProAsnLysLeuTrp 105  
 Db 1494 TGGAAACACCTCCGACTATGCGGAGTTAAGGATCTGCGAATACCGCGCATCGCATCTGG 1553  
 QY 106 LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrThrGlnThr 125  
 Db 1554 AAGCCGGACGTGCTGATGTACAAACAGTCGGAGTTGAGGGATTTGACGACCTACCGACG 1613  
 QY 126 AsnValValValArgSerGlyGlySerCysLeuTyrValProProGlyIlePheLysSer 145  
 Db 1614 AAGTGGTGGTGGGAACACGCTCGTGTCTATAGTTCGCGCGGGGATCTTCAAGTCG 1673  
 QY 146 ThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMetLysPhe 165  
 Db 1674 ACGTCAAGATCGACATCAGTGGTCCCTTCGATGACCGCGGTGCGAGATGAAGTTC 1733  
 QY 166 GlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyGly 185  
 Db 1734 GGCAGTTGGACCTACGACGATTTCCAGCTGGATTTACAATTAAGATGAACCTGGCGGT 1793  
 QY 186 AspLeuSerAspPheIleThrAsnGlyGluTyrTrpLeuIleGlyMetProGlyLysLys 205  
 Db 1794 GATATCAGCAGTTACGTGCTCAACGGCGAGTGGAACTACTGGGTGTCGCCGCAACGT 1853  
 QY 206 AsnThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMet 225  
 Db 1854 AACGAGATCTATTACAACCTGCTGCCCGAACCCCTATATAGACATCACCTTCGCCATCATC 1913  
 QY 226 IleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleSer 245  
 Db 1914 ATCCGCCGCGAACACTGTACTATTCTTCAACCTGATCATCACTTGTGTACTGTGCTGCC 1973  
 QY 246 SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuThrLeuGly 265  
 Db 1974 TCCATGGCCCTTGTCTGGATTCACCTGCGCGCAGATTCGGGTGAAATAATTTCGCGGGT 2033  
 QY 266 ValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGln 285  
 Db 2034 GTTACCATCTTGTCTCGTGGACCGGTGTTCTGAATATGTGTTGCCGAGACAATGCCGGCT 2093  
 QY 286 ValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAla 305



QY 184 GlyGlyAspLeuSerAspPheLeuThrAsnGlyGluThrPheLeuLeuGlyMetProGly 203  
 Db 598 GGTGGCGACATCTTACGCTTTATACCAATGGCAATGGGACTTGTAGGTGTGCCCGGT 657  
 QY 204 LysLysAsnThrLeuThrTyrAlaCysCysProGluProTyrValAspValThrPheThr 223  
 Db 658 AACGAAATGAATCTACTATAATGTGTCGCCAAGCTTATATTGACATAACATTCGCC 717  
 QY 224 IleMetIleArgArgArgThrLeuTyrTyrPhePheAsnLeuLeuValProCysValLeu 243  
 Db 718 ATTTGTGTAAGGCGCAAAACGTTGCTACTATTTTCAATCTGATTGTGCGGTGCTACTG 777  
 QY 244 IleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLeuThr 263  
 Db 778 ATCCCTTCATGGCACTGTAGGGTTTACAGTCCACAGATCTGTGTGAAGCTTTGCG 837  
 QY 264 LeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeu 283  
 Db 838 CTGGAGTTACATCTTATTATCGTTACAGTCTTCTCAACATGGTGGCGGAACAATG 897  
 QY 284 ProGlnValSerAspAlaIleProLeuLeuGly 294  
 Db 898 CCGCGCACTCCGATCGGTACCGCTGCTCGGT 930

## RESULT 6

AAC58395  
 ID AAC58395 standard; cDNA; 1509 BP.  
 AC AAC58395;  
 XX  
 XX 29-JAN-2001 (first entry)  
 XX Human PRO2145 nucleotide sequence SEQ ID NO:76.  
 XX Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth;  
 KW proliferation; tumorigenesis; identification; cancer; cytostatic;  
 KW neurotropic; neuroprotective; antiinflammatory; immunosuppressive;  
 KW immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy;  
 KW neuronal disorder; glial disorder; astrocytal disorder; angiogenic;  
 KW hypothalamic disorder; glandular disorder; macrophagal disorder;  
 KW epithelial disorder; stromal disorder; blastocoelec disorder;  
 KW inflammatory disorder; immunologic disorder; ss.  
 XX Homo sapiens.  
 OS  
 XX WO200053755-A2.  
 FN  
 XX 14-SEP-2000.  
 PD  
 XX 06-JAN-2000; 2000WO-US000376.  
 PF  
 XX 08-MAR-1999; 99WO-US0005028.  
 PR 02-JUN-1999; 99WO-US012252.  
 PR 23-JUN-1999; 99US-0141037P.  
 PR 07-JUL-1999; 99US-0143048P.  
 PR 26-JUL-1999; 99US-0145698P.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 20-DEC-1999; 99WO-US030911.  
 PR 05-JAN-2000; 2000WO-US000219.  
 XX (GETH ) GENENTECH INC.  
 PA  
 XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA;  
 PI Watanabe CK, Wood WI;  
 PI WPI; 2000-572270/53.  
 DR P-PSDB; AAB24088.  
 DR  
 XX Thirty PRO polynucleotides encoding PRO polypeptides, useful in the  
 PT treatment, diagnosis and prevention of cancer.  
 PT  
 XX Claim 50; Fig 57; 286pp; English.  
 PS

XX

The present invention describes an isolated antibody that binds to one of the human PRO proteins designated PRO212, PRO290, PRO341, PRO535, PRO619, PRO717, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009, PRO1025, PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184, PRO1187, PRO1281, PRO1293, PRO1309, PRO834, PRO1317, PRO1710, PRO2094, PRO2145 OR PRO2198. PRO antagonists can be used to inhibit tumour cell growth. The PRO polypeptides and nucleotides are useful in the treatment, diagnosis and prevention of cancer. The antibodies and other anti-tumour compounds may be used to treat various conditions, including those characterised by overexpression and/or activation of the amplified PRO genes. Exemplary conditions or disorders to be treated with such antibodies and other compounds include benign or malignant tumours (e.g., renal, liver, kidney, bladder, breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulva, thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and neck tumours), leukaemias and lymphoid malignancies, other disorders such as neuronal, glial, astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal and blastocoelec disorders, and inflammatory, angiogenic and immunologic disorders. AAC58242 to AAC58366 represent PCR primers and hybridisation probes used in the isolation of the human PRO sequences. AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human PRO polynucleotide and protein sequences given in the exemplification of the present invention

SQ Sequence 1509 BP; 297 A; 451 C; 429 G; 332 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1-46e-124 Length: 1509  
 Score: 1258.50 Matches: 250  
 Percent Similarity: 63.23% Conservative: 75  
 Best Local Similarity: 48.64% Mismatches: 138  
 Query Match: 47.67% Indels: 51  
 DB: 3 Gaps: 7

US-09-303-232-6 (1-501) x AAC58395 (1-1509)

QY 8 LeuAlaLeuLeuAla---LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26  
 Db 28 CTGGCGCTGGCGCGTCTCTCTCTGCGTCTCTCTGCGGCGAGTTCAGAGAGAGCTT 87  
 QY 27 LeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSerGlu 46  
 Db 88 TACAAGGAGCTGGTCAAGAACTACAATCCCTTGGAGAGGCCCGTGGCAATGATCTGCCAA 147  
 QY 47 ProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleIleAspValAspGluLysAsn 66  
 Db 148 CCACTCACCGTCTACTTCTCCCTGAGCCCTCTCGAGATCATGCGTGGATGAGAGAAC 207  
 QY 67 GlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArgTrp 86  
 Db 208 CAAGTTTAAACCAACCAACATTTGGCTGCAAAATCTCTTGGACAGATCACTATTACAGTGG 267  
 QY 87 AsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThrProAsnLysLeuTrpLys 106  
 Db 268 AATGTGTCAGAAATATCCAGGGGTGAAGCTGTTCGTTTCCAGATGCCAGATTTGGAAA 327  
 QY 107 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126  
 Db 328 CCAGACATTTCTCTATACAGTGTGTGAGAGCGCTTTGACGCCACATCTCCACTAAC 387  
 QY 127 ValValValArgSerGlyGlySerCysLeuTyrValProGlyIlePheLysSerThr 146  
 Db 388 GTGTTGGTGAATTTCTTCTGGCATTTGCCAGTACCTGCTCCAGGCATATTCAAGAGTTC 447  
 QY 147 CysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMetLysPheGly 166  
 Db 448 TGTACATCGATGTACGGTGGTTCCCTTTGATGTGCAGACTGCAACTGAAGTTTGGG 507  
 QY 167 SerThrThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyGlyAsp 186  
 Db 508 TCTGTGCTTACGGAGGCTGGTCTCTGGATCTGCAGATCGAGAG-----GCAGAT 558



Db 328 CCAGACATCTCTCTATACAGTGTCTGATGAGCGCTTTGACGCCACATTCACACATAAC 387  
Qy 127 ValValValArgSerGlyGlySerCysLeuTyrValProProGlyIlePhePheSerThr 146  
Db 388 GTGTGTGTGAATCTCTGGGCATTTGCCAGTACCTGCTCCAGGCATATTCAGAGTTCC 447  
Qy 147 CysLysMetAspIleAlaTyrPheProPheAspGlnHisCysAspMetLysPheGly 166  
Db 448 TGCTACATCATGTACGCTGTGTTCCTTTGATGTGCAGCACTCAAACTGAATGTGG 507  
Qy 167 SerTyrThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyGlyAsp 186  
Db 508 TCCTGGTCTTACGGAGGCTGTCTCTGATCTGCAGATGCAGAG-----GCAGAT 558  
Qy 187 LeuSerAspPheIleThrAsnGlyGluTyrTyrLeuIleGlyMetProGlyLysAsn 206  
Db 559 ATCAGTGGCTATATCCCAATGGAGATGGACCTAGTGGGAATCCCGGCAAGAGGAGT 618  
Qy 207 ThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMetIle 226  
Db 619 GAAAGGTTCTATGAGTGTGCAAGAGCCCTACCCGATGTCACTTCACAGTACCATG 678  
Qy 227 ArgArgArgThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleSerSer 246  
Db 679 CGCGCAGAGCGCTCTACTATGCGCTCAACCTGCTGATCCCTGTGTCTCATCTCCGCC 738  
Qy 247 MetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuThrLeuGlyVal 266  
Db 739 CTGCGCTCTGTGTGTCTCTCTCTGCGATTCGCGGAGAAATTCCTCTGGGATA 798  
Qy 267 ThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnVal 286  
Db 799 ACAGTCTTACTCTCTCTACGCTCTTATGCTGTCTGTGCTGATCATGCCCGCAACA 858  
Qy 287 SerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSer 306  
Db 859 TCCGATTCGGTACCATTCATGACCCAGTACTTCCGACGACCATGATCATCTGGGGCTC 918  
Qy 307 SerValValLeuThrValValValLeuAsnTyrHisHsargThrAlaAspIleHisGlu 326  
Db 919 TCGTGTGTGTGAGGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 978  
Qy 327 MetProGlnThrIleLysSerValPheLeuGlnThrLeuProTyrIleLeuArgMetSer 346  
Db 979 ATGCCAAGTGGACGACGATCTCTCTGAAGTGTGCGGTGCTGCTGCTGCTGCTGCTG 1038  
Qy 347 ArgProGlyLysIleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeu 366  
Db 1039 AGGCCCGGGAGGACCAAGGTGCGCGCGCTGCGCAGCAGCAGCGCGCTGCGAGCTG 1098  
Qy 367 GluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAsp 386  
Db 1099 GCCAGTGTGAGATGACGCGCGTGGCG-----1125  
Qy 387 PheArgHisGlyProProProAsnSerThrAlaSerThrGlyAsnLeu-----403  
Db 1126 -----CGCGCGCC-----GCCAGCAACGGGAACCTGCTGTATC 1161  
Qy 404 -----GlyProGlyCys 407  
Db 1162 GGCTTCGCGCGCTGGACGCGTGCATGTGTCGCGACCCCGCTGCGGGTAGTGTGT 1221  
Qy 408 SerIlePheArgThrAspPheArgArgSerPheValArgProSerThrMetGluAspVal 427  
Db 1222 GGC-----CGCATGCGCTGTCTCCCGCCACGACGATGACGACCTC 1260  
Qy 428 -----GlyGlyGlyLeuGlySerHisArgGluLeuHisLeuIleLeuArgGluLeu 445  
Db 1261 CTGCACGGCGGCAACCCCGAGGGGACCGGATTCGCAAGATCTCGAGGAGGTC 1320  
Qy 446 GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuSerAspTyr 465  
Db 1321 CGTACATTCCTCAATCGCTTCCGCTGCGCAGGACGAAAGCGAGGCGGTCTCAGCGAGTGG 1380

Qy 466 LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle 485  
Db 1381 AAGTTCCGCGCTGTGTGTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
Qy 486 IleAlaThrValAlaValLeuLeuSerAlaProHisIleIle 499  
Db 1441 ATCTGCACCATCGCATCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1482

RESULT 8  
AAV12197  
ID AAV12197 standard; cDNA; 1876 BP.  
AC AAV12197;  
XX 14-MAY-1998 (first entry)  
XX Human neuronal nicotinic acetylcholine receptor alpha-7 subunit cDNA.  
XX Human; neuronal nicotinic acetylcholine receptor; alpha-7 subunit;  
KW brain tissue; screening; NACHR; antibody; ds.  
XX Homo sapiens.  
FH Location/Qualifiers  
FT 73..1581  
FT /\*tag= a  
FT /product= "neuronal nicotinic acetylcholine receptor  
FT alpha-7 subunit"

WO9420617-A2.

15-SEP-1994.

08-MAR-1994; 94WO-US002447.

08-MAR-1993; 93US-00028031.

(SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.

Elliott KJ, Ellis SB, Harpold MM;

WPI; 1994-303024/37.

P-PSDB; AAM44153.

Human neuronal nicotinic acetylcholine receptor subunits and DNA - also transformed cells useful for screening cpds. which modulate activity of the receptor.

Claim 8; Page 78-79; 99pp; English.

The present sequence encodes a human neuronal nicotinic acetylcholine receptor (NACHR) subunit. The cells expressing the alpha and/or beta NACHR subunits may be used in a method of screening compounds to identify any which modulate the activity of human neuronal NACHR. Subunit specific antibodies may be used to monitor the distribution and expression density of various subunits in normal vs diseased brain tissues. Testing of single receptor subunits or specific receptor subunit combinations with a variety of potential agonists or antagonists provides information with respect to the function and activity of the individual subunits and should lead to the identification and design of compounds that are capable of very specific interaction with one or more receptor subtypes. The resulting drugs should exhibit fewer unwanted side effects than drugs identified e.g. screening with cells that express a variety of subtypes

SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	2,02e-124	Length:	1876
Score:	1258.50	Matches:	250
Percent Similarity:	63.23%	Conservative:	75
Best Local Similarity:	48.64%	Mismatches:	138
Query Match:	47.67%	Indels:	51

```

DB:      2          Gaps:      7
US-09-303-232-6 (1-501) x AAV12197 (1-1876)
Qy      8 LeuAlaLeuLeuAla---LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26
      100 CTGGCGTGGCGGTCGCTCTCGACGTGTCCCTGCAAGCGAGTTCACAGAGGAAGCTT 159
      27 LeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSerGlu 46
      160 TACAAGAGAGTGGTGCAGAACTACAATCCCTTTGAGAGGCGCCGTGGCCAAATGACTCCAA 219
      47 ProLeuGluValArgPheGlyLeuThrLeuGlnIleLeuAspValAspGluLysAsn 66
      220 CCATCTACCGTCTACTTCTCCCTGAGCTCTCTGAGATCATGGAGCTGGATGAGAAGAC 279
      67 GlnLeuLeuLeuThrAsnIleThrLeuSerLeuThrAsnGluThrAsnLeuArgTyr 86
      280 CAAGTTTAAACCAACCAATTTGGCTGCAATGTCTTGGACAGATCATATTATACAGTGG 339
      87 AsnAspSerGluTyrGlyValLysAspLeuArgIleThrProAsnLysLeuThrLys 106
      340 AATGTGTGAGAATATCCAGGGGTGAAGACTTTCGTTTCCAGATGGCCAGATTGGAAA 399
      107 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126
      400 CCAGACATCTCTCTATAACAGTGTCTGATGAGCGCTTTGACGCCACATCCACACTAAC 459
      127 ValValValArgSerGlyGlySerCysLeuTyrValProGlyIlePheLysSerThr 146
      460 GTGTTGGTGAATCTCTCGGCATTTGCCAGTACCTGCTCCAGGCATATTCAAGAGTTCC 519
      147 CysLysMetAspIleAlaThrPheProPheAspAspGlnHisCysAspMetLysPheGly 166
      520 TGCTACATCGATGACGTGTGTTTCCCTTTGATGTCAGCACTGCAAACTGAAGTTTGGG 579
      167 SerTyrThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyLys 186
      580 TCCTGGTCTTACGGAGGCTGCTCTTGGATCTGCAGATGCAGGAG-----GCAGAT 630
      187 LeuSerAspPheIleThrAsnGlyGluTyrTyrLeuIleGlyMetProGlyLysLysAsn 206
      631 ATCAGTGGCTATATCCCAATGGAGATGGGACCTAGTGGGAATCCCGGCAAGAGAGT 690
      207 ThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMetIle 226
      691 GAAAGTTCTATGAGTGTCTGAAGAGCCCTACCCCGATGACCTTCACAGTGACCAATG 750
      227 ArgArgArgThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleSerSer 246
      751 CGCCGACGACGCTCTACTATGGCTCAACCTGTGTGATCCCTGTGTGCTCATCTCCGCC 810
      247 MetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuThrLeuGlyVal 266
      811 CTCGCCCTGTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 870
      267 ThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnVal 286
      871 ACAGTCTTACTCTCTCTACCGTCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
      287 SerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSer 306
      931 TCCGATTCCGTACCAATGATGATCCAGTACTTCGCCAGCACCATCATCATCTGCGGCCTC 990
      307 SerValValLeuThrValValValLeuAsnTyrHisArgThrAlaAspIleHisGlu 326
      991 TCGGTGGTGGTACCGGTGATGCTGCTGAGTACCAACACACAGCCCGGAGGGGGCAG 1050
      327 MetProGlnTyrIleLysSerValPheLeuGlnTyrLeuProTyrIleLeuArgMetSer 346
      1051 ATGCCCAAGTGACACAGATCATCTCTGAACTGCTGGCGGTGCTCTCTGATGATGAAG 1110
      347 ArgProGlyLysLysIleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeu 366

```

```

      1111 AGCCCGGGGAGGACAAAGTGTGCGCCCGCCCTGCGACGACAAAGCGCGCTGACGCGTG 1170
      367 GluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAsp 386
      1171 GCCAGTGTGAGATGAGCGCGCTGGCG----- 1197
      387 PheArgHisGlyProProProAsnSerThrAlaSerThrGlyAsnLeu----- 403
      1198 -----CCGCGGCC-----GCCAGCAACGGGAACCTGCTGTACATC 1233
      404 -----GlyProGlyCys 407
      1234 GGCTTCGCGCGCTGAGCGGCTGCACTGTGTCCGACCCCGGACTCTGGGGTAGTGTGT 1293
      408 SerIlePheArgThrAspPheArgSerPheValArgProSerThrMetGluAspVal 427
      1294 GGC-----CGCATGGCGCTGCTCCCGGACTTGGCCAAAGATCTCTGGAGAGGTC 1392
      428 -----GlyGlyGlyLeuGlySerHisHisArgGluLeuHisLeuLeuArgGluLeu 445
      1333 CTGCACGCGGGCAACCCCGAGGGGACCCCGGACTTGGCCAAAGATCTCTGGAGAGGTC 1392
      446 GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuIleSerAspTyr 465
      1393 CGCTTACATTGCCAATCGCTTCCGCTGCCAGGACGAAAGCGAGCGGCTCTGCAGCGAGTGG 1452
      466 LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle 485
      1453 AAGTTCGCGCGCTGTGTGGAGCCGCTGTGCCTCATGGCGCTTCTCGGTCTTACCATC 1512
      486 IleAlaThrValAlaValLeuLeuSerAlaProHisIleIle 499
      1513 ATCTGCACCATCGCATCTGATGTGCGCTCCCAACTTCGCTG 1554

RESULT 9
AAT48239
ID AAT48239 standard; DNA; 1876 BP.
AC AAT48239;
XX 09-APR-1997 (first entry)
DT 09-APR-1997 (first entry)
DE Neuronal nicotinic acetylcholine receptor alpha-7 subunit DNA.
XX Neuronal nicotinic acetylcholine receptor; nAChR; neurotransmitter;
KW ligand-gated receptor; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 73..1581
XX /*tag= a
XX
XX WO9641876-A1.
XX
XX 27-DEC-1996.
XX 07-JUN-1996; 96WO-US009775.
XX 07-JUN-1995; 95US-00484722.
XX (SIBI-) SIBIA NEUROSCIENCES INC.
XX Elliott KJ, Harpold WM;
XX WPI; 1997-065463/06.
XX P-PSDB; AAW09025.
XX Nucleic acids encoding nicotinic acetylcholine receptor sub-units - used
XX in screening to determine the effect of drugs on the receptor.
XX Disclosure; Page 71-73; 108pp; English.

```



XX A DNA sequence (AAT48239) codes for the alpha-7 subunit (AAW09025) of the  
 CC human neuronal nicotinic acetylcholine receptor (nAChR). Host cells, esp.  
 CC mammalian cells or amphibian oocytes, carrying alpha-7 nucleic acids,  
 CC opt. in combination with other alpha and/or beta subunit nucleic acids  
 CC (see also AAT48232-38, AAT48240-41), express recombinant nAChR subunits  
 CC useful for identifying cpds. that modulate the activity of human nAChRs  
 XX  
 SQ Sequence 1876 BP; 369 A; 553 G; 531 C; 423 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,02e-124 Length: 1876  
 Score: 1258.50 Matches: 250  
 Percent Similarity: 63.23% Conservative: 75  
 Best Local Similarity: 48.64% Mismatches: 138  
 Query Match: 47.67% Indels: 51  
 DB: 2 Gaps: 7

US-09-303-232-6 (1-501) x AAT48239 (1-1876)

Qy 8 LeuAlaLeuLeuAla---LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26  
 Db 100 CTGGCGCTGGCGCGTCTCTCTGACGTGCTCCGCAAGCGCGAGTTCCAGAGGAAGCTT 159  
 Qy 27 LeuAsnAlaLeuLeuAlaAsnThrLeuGluArgProValAlaAsnGluSerGlu 46  
 Db 160 TACAAGGAGCTGGTCAAGAACAATACTCCCTTGGAGAGCGCGTGGCAATGATCTGCCAA 219  
 Qy 47 ProLeuGluValArgPheGlyLeuThrLeuGlnGlnLeuLeuLeuLeuLeuLeuLeu 66  
 Db 220 CCACTCCCGTCTACTTCTCCCTGAGCTCTGAGGATCATGAGCGTGGATGAGAGAAC 279  
 Qy 67 GlnLeuLeuLeuThrAsnLeuThrLeuSerLeuGluThrAsnAspThrLeuLeuArgTrp 86  
 Db 280 CAAGTTTAAACCAACCAATTTGGCTGCAATGCTTGGACAGATCACTATTACAGTGG 339  
 Qy 87 AsnAspSerGluTrpGlyGlyValLysAspLeuArgGlnLeuThrProAsnLysLeuTrpLys 106  
 Db 340 AATGTGTCAGAAATATCCAGGGGTGAGACTGTTCTGTTCCAGATGGCCAGATTGGAAA 399  
 Qy 107 ProAspValLeuMetThrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126  
 Db 400 CCAGACATCTCTCTATACAGTGTGATGAGCGCTTTGACGCCACATTCACACATAAC 459  
 Qy 127 ValValValArgSerGlyGlySerCysLeuThrValProProGlyIlePheLysSerThr 146  
 Db 460 GTGTGTGTGAATCTCTGGGCATTTGGCAGTACCTCCGCTCCAGGCATATTCAAGAGTTC 519  
 Qy 147 CysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMetLysPheGly 166  
 Db 520 TGCTACATCGATGTACGCTGGTTCCTTTGATGTGAGCACTGCAAACTGAAGTTTGGG 579  
 Qy 167 SerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLysAspGluAlaGlyGlyAsp 186  
 Db 580 TCCTGTGCTTACGGAGCGTGTCTCTGATCTGAGATGCGAGAG-----GCAGAT 630  
 Qy 187 LeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGlyMetProGlyLysLysAsn 206  
 Db 631 ATCAGTGGCTATATCCCAAGAGATGGAGCTAGTGGGAATCCCGGCAAGAGAGT 690  
 Qy 207 ThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMetIle 226  
 Db 691 GAAAGGTCTATGATGCTGCAAGAGCCCTACCCGATGTCCACTTCACAGTACCATC 750  
 Qy 227 ArgArgArgThrLeuTyrTyrPhePheAsnLeuLeuValProCysValLeuLeuSerSer 246  
 Db 751 CGCGCAGGACGCTCTACTATGCGCTCAACCTGCTGATCCCTGTGTCTCATCTCCGCC 810  
 Qy 247 MetAlaLeuGlyPheThrLeuProProAspSerGlyGluLysLeuThrLeuGlyVal 266  
 Db 811 CTCGCCCTGTGGTGTCTCTGCTGCTCCGATTCGGGAGAGATTCCTCCCTGGGATA 870  
 Qy 267 ThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnVal 286

Db 871 ACAGTCTACTCTCTCTTACCGTCTTATGCTGTCTGGGTGTAGATCATGCCGCCAACA 930  
 Qy 287 SerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSer 306  
 Db 931 TCCGATTTCGTAACATTGATAGCCAGTACTTCGCCAGCACCATGATCATCTGTGGCCCTC 990  
 Qy 307 SerValValLeuThrValValValLeuAsnTyrHisArgThrAlaAspIleHisGlu 326  
 Db 991 TCGGTGTGTGTGACGGTGATCGTGTGCTGACCTACACACACACCGCGCGGCGCAAG 1050  
 Qy 327 MetProGlnTrpIleLysSerValPheLeuGlnTrpLeuProTrpIleLeuArgMetSer 346  
 Db 1051 ATGCCCAAGTGGACACGAGATCATCTTCTGAACCTGGTGGCGGTCTCTCGGAATGAAG 1110  
 Qy 347 ArgProGlyLysLysIleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeu 366  
 Db 1111 AGGCCCGGGAGGACAAAGTGTGCGCGCGCTGCCAGACACAGCGCGGTGCAGCCTG 1170  
 Qy 367 GluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAsp 386  
 Db 1171 GCCAGTGTGAGATGAGCGCGCTGGCG-----1197  
 Qy 387 PheArgHisGlyProProProAsnSerThrAlaSerThrGlyAsnLeu-----403  
 Db 1198 -----CCGCCGCC-----GCCAGCAACGGGAACCTCTCTGTACATC 1233  
 Qy 404 -----GlyProGlyCys 407  
 Db 1234 GGCTTCCGCGCGCTGGACGGCGTCACTGTGTCCGACCCCGACTCTGGGGTAGTGTG 1293  
 Qy 408 SerIlePheArgThrAspPheArgSerPheValArgProSerThrMetGluAspVal 427  
 Db 1294 GGC-----CGCATGCGCTGTCTCCCGCACGACGATGAGCACCTC 1332  
 Qy 428 -----GlyGlyGlyLeuGlySerHisArgGluLeuHisLeuLeuLeuArgGluLeu 445  
 Db 1333 CTGCACGCGCGGCAACCCCGAGGGGACCCGACTTGGCCAGATCTCTGGAGAGGTC 1392  
 Qy 446 GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuLeuSerAspTrp 465  
 Db 1393 CGCTACATTGCCAATCGTCTCCGCTGCGAGACGAAAGCGAGCGGTCTGCAGCGAGTGG 1452  
 Qy 466 LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle 485  
 Db 1453 AAGTTCGCCCGCTGTGTGTGGACCGCTGTGCTCATGGCTTCTCGTCTTCCCATC 1512  
 Qy 486 IleAlaThrValAlaValLeuLeuSerAlaProHisIleIle 499  
 Db 1513 ATCTGCACCATCGGCATCTTGTGCTGGCTCCCAACTTCGTG 1554  
 RESULT 10  
 ABS54875  
 ID ABS54875 standard; cDNA; 1876 BP.  
 AC ABS54875;  
 XX  
 XX 06-DEC-2002 (first entry)  
 DT  
 XX Human neuronal nicotinic acetylcholine receptor alpha 7 subunit cDNA.  
 DE  
 XX Human; neuronal nicotinic acetylcholine receptor; nNACHR; gene; ss;  
 KW ion flux; alpha 7 subunit.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 PH 73..1581  
 FT /\*tag= a  
 FT /product= "Human nNACHR alpha 7 subunit"  
 XX  
 PN US6440681-B1.  
 XX

PD 27-AUG-2002.  
 XX 07-JUN-1995; 95US-00487596.  
 PF 03-APR-1990; 90US-00504455.  
 PR 30-NOV-1992; 92US-00938154.  
 PR 08-MAR-1993; 93US-00028031.  
 PR 08-NOV-1993; 93US-00149503.  
 XX (MERI ) MERCK & CO INC.  
 FA Elliott KJ, Ellis SB, Harpold MW;  
 XX WPI; 2002-711528/77.  
 XX P-PSDB; ABG70492.  
 DR Identifying antagonists or agonists of human neuronal nicotinic  
 PT acetylcholine receptors, by contacting recombinant cells with test  
 PT compound, and measuring ion flux of cells or binding of compound to  
 PT nNACHr.  
 XX Claim 101; Col 57-60; 56pp; English.  
 XX The invention relates to a method for identifying compounds that are  
 CC antagonists or agonists of human neuronal nicotinic acetylcholine  
 CC receptors (nNACHrs), by contacting recombinant cells with a test compound  
 CC and measuring ion flux, the electrophysiological response of the cells or  
 CC binding of the test compound to the nNACHr. The recombinant cells are  
 CC produced by transfection with a nucleic acid encoding at least one human  
 CC nNACHr (alpha or beta) subunit, such that the cells express an nNACHr  
 CC comprising one human subunit encoded by the transfected nucleic acid.  
 CC This sequence represents cDNA encoding the alpha 7 subunit of the human  
 CC nNACHr polypeptide  
 XX SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Alignment Scores:  
 Pred. No.: 2,02e-124 Length: 1876  
 Score: 1258.50 Matches: 250  
 Percent Similarity: 63.23% Conservative: 75  
 Best Local Similarity: 48.64% Mismatches: 138  
 Query Match: 47.67% Indels: 51  
 DB: 6 Gaps: 7

US-09-303-232-6 (1-501) x ABS54875 (1-1876)

QY 8 LeuAlaLeuAla---LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26  
 DB 100 CTGGCGCTGGCGCGCTGCTCTCCACGCTGCTCCGCAAGCGAGTTCAGAGGAAGCTT 159  
 QY 27 LeuAsnAlaLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSerGlu 46  
 DB 160 TACAGGAGCTGGTCAAGAACTACATCCCTTGGAGAGCGCGTGGCCCAATGACTCGCAA 219  
 QY 47 ProLeuGluValArgpGlyLeuThrLeuGlnGlnIleLeuAspValAspGluLysAsn 66  
 DB 220 CCACTCACCGTCTACTCTCCCTGAGCTCTCTGAGAGCGCGTGGCCCAATGACTCGCAA 279  
 QY 67 GlnLeuLeuIleThrAsnIleThrLeuSerLeuGluTrpAsnAspTyrAsnLeuArgTrp 86  
 DB 280 CAAGTTTAAACACCAACATTTGGCTGCAATGTCTTGGACAGATCATGTGATGAGAAGAAC 339  
 QY 87 AsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThrProAsnLysLeuTrpLys 106  
 DB 340 AATGTGTGAGAATATCCAGGGGTGAAGACTGTCTGTTCCAGATGGCCAGATTTGGAAA 399  
 QY 107 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126  
 DB 400 CCAGACATTTCTCTATAACAGTGTCTGATGAGCGCTTTGAGCCACACATTCACACTAAC 459  
 QY 127 ValValValArgSerGlyGlySerCysLeuTyrValProGlyIlePheLysSerThr 146  
 DB 460 GTGTTGGTGAATTCTTCTGGGCAATGCCAGTACCTGCTCCAGGCATATTCAGAGTTCC 519

QY 147 CysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMetLysPheGly 166  
 DB 520 TGCTACATCGATGATCGCTGGTTTCCCTTTGATGTGAGCACTGCAAACTGAAGTTGGG 579  
 QY 167 SerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyGlyAsp 186  
 DB 580 TCCTGGTCTTACGAGGCTGGTCTTGATCTGCAGATGCAGGAG-----GCAGAT 630  
 QY 187 LeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGlyMetProGlyLysLysAsn 206  
 DB 631 ATCAGTGGCTATATCCCAATGGGAATGGGACCTAGTGGGAATCCCGGCAAGAGGAGT 690  
 QY 207 ThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheIleMetIle 226  
 DB 691 GAAAGGTTCTATGAGTGTCTCAAGAGCCCTACCCCGATGTCACCTTCACATGACCATG 750  
 QY 227 ArgArgArgThrLeuTyrTyrPheAsnLeuIleValProCysValLeuIleSerSer 246  
 DB 751 CGCGCAGGAGCGCTTACTATGGCTCAACCTGTGTATCCCTGTGTCTCATCTCCGCC 810  
 QY 247 MetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuThrLeuGlyVal 266  
 DB 811 CTCGCCCTGCTGGTGTCTCTCTCGAGATTCGGGGAGAAATTCCTCTGGGAAT 870  
 QY 267 ThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnVal 286  
 DB 871 ACAGCTTACTCTCTTACCGTCTTCATGCTGTCTGTGGTGTGAGATCATGCCCAACA 930  
 QY 287 SerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSer 306  
 DB 931 TCCGATTGCGTACCATTTGATAGCCAGTACTTCCGCCAGCATGATCATCTGGGGCTC 990  
 QY 307 SerValValLeuThrValValLeuAsnTyrHisArgThrAlaAspIleHisGlu 326  
 DB 991 TCGGTGGTGGTGGAGTGTGCTGAGTACCATCCACACAGCCCGCCGCGGGGCAAG 1050  
 QY 327 MetProGlnTrpIleLysSerValPheLeuGlnTrpLeuProTrpIleLeuArgMetSer 346  
 DB 1051 ATGCCCAAGTGACCAAGTATCTCTTGAACCTGGTGGCGTGTCTCTCGAATGAAG 1110  
 QY 347 ArgProGlyLysLysIleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeu 366  
 DB 1111 AGGCCCGGGAGGACAAGGTGCGCCGCGCTCGCCAGCACAGCAGCGCGTGCAGCCTG 1170  
 QY 367 GluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAsp 386  
 DB 1171 GCCAGTGTGGAGATGAGCGCGCTGGCG-----GlyProGlyCys 407  
 QY 387 PheArgHisGlyProProProProAsnSerThrAlaSerThrGlyAsnLeu----- 403  
 DB 1198 -----CCGCCGCC-----GCCAGCAACGGGAACCTGTGTGTACATC 1233  
 QY 404 -----GlyProGlyCys 407  
 DB 1234 GGCTTCGCGCGCTGGACGGGTGCTGTGTCTCCGACCCCGCATCTCTGGGTAGTGTGT 1293  
 QY 408 SerIlePheArgThrAspPheArgSerPheValArgProSerThrMetGluAspVal 427  
 DB 1294 GGC-----CGCATGGCTGTCTCCCGACGACGACGATGAGCACCTC 1332  
 QY 428 -----GlyGlyGlyLeuGlySerHisArgGluLeuHisLeuLeuArgGluLeu 445  
 DB 1333 CTGCACGCGGGGCAACCCCGAGGGGACCCCGACTTGGCAAGATCCTCGAGGAGGTC 1392  
 QY 446 GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuIleSerAspTrp 465  
 DB 1393 CGCTACATTTGCCAATCGCTTCCGTCCGACGAGCAAAACGAGGCGGTCTGCAGCGAGTGG 1452  
 QY 466 LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle 485  
 DB 1453 AAGTTCCGCGCTGTGTGTGGACCGCTGTGCCTCATGGCTTCTCGTCTTCACCATC 1512



```

Db 1171 GCCAGTGTGGAGATGAGCGCGTGGCG----- 1197
QY 387 PheArgHisGlyProProProAsnSerThrAlaSerThrGlyAsnLeu----- 403
Db 1198 -----CGCGCGCC-----GCCAGCAACGGAACCTGCTGTACATC 1233
QY 404 -----GlyProGlyCys 407
Db 1234 GGCTTCGCGCGCTGGAGCGGTGCACCTGTCTCCGACCCCGACTCTGGGGTAGTGTGT 1293
QY 408 SerIlePheArgThrAspPheArgSerPheValArgProSerThrMetGluAspVal 427
Db 1294 GGC-----CGCATGCGCTGCTCCCGCAGCAGCAGATGAGCAGCTC 1332
QY 428 -----GlyGlyGlyLeuGlySerHisArgGluLeuHisLeuLeuLeuArgGluLeu 445
Db 1333 CTGCACGCGCGGCAACCCCGAGGGGACCCGAGACTTGGCCAAAGATCTCTGGAGAGTCT 1392
QY 446 GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuLeuSerAspTrp 465
Db 1393 CGCTACATTCGAATCGCTTCGCTGCCAGGACGAGAGCGGCTCTGCAGCGAGTGG 1452
QY 466 LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle 485
Db 1453 AAGTTCGCGCGCTGTGTGTGGACCGCGCTGTGCCTCATGGGCTTCTCGTCTTCCACCATC 1512
QY 486 IleAlaThrValAlaValLeuLeuSerAlaProHisIleIle 499
Db 1513 ATTCGACCATCGCATCTGATGTCGCTCCCAACTTCGTG 1554

RESULT 12
ID ADA10864
XX ADA10864 standard; DNA; 1876 BP.
AC ADA10864;
XX
XX 06-NOV-2003 (first entry)
XX
XX Human neuronal nicotinic acetylcholine receptor alpha 7 subunit DNA.
KW ds; gene; alpha 7 subunit; human;
KW neuronal nicotinic acetylcholine receptor; ligand-gated ion channel;
KW synaptic transmission; gene therapy; transgenic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX 5'UTR 1..72
XX /*tag= a
XX CDS 73..1584
XX /*tag= b
XX /*product= "nAChR alpha 7 subunit"
XX 3'UTR 1585..1876
XX /*tag= c
XX
XX US6524789-B1.
XX
XX 25-FEB-2003.
XX
XX 07-JUN-1996; 96US-00660451.
XX
XX 07-JUN-1995; 95US-00484722.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Elliott KJ, Harpold MM;
XX
XX WPI; 2003-511917/48.
XX P-PSDB; ADA10874.
XX
XX New nucleic acid encoding an alpha-6 or a beta-3 subunit of a human
XX neuronal nicotinic acetylcholine receptor (nAChR), useful for identifying
XX compounds that modulate human neuronal nAChR activity.

```

```

XX Claim 33; Col 67-72; 63pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule comprising a
XX nucleotide sequence encoding an alpha 6 or a beta 3 subunit of a human
XX neuronal nicotinic acetylcholine receptor nAChR. nAChR's form ligand-
XX gated ion channels that mediate synaptic transmissions between nerve and
XX muscle and between neurons upon interaction with the neurotransmitter
XX acetylcholine. The nucleic acid molecule is useful for identifying
XX compounds that modulate human neuronal nAChR. The present sequence
XX represents DNA encoding the human neuronal nicotinic acetylcholine
XX receptor, nAChR, alpha 7 subunit. Note: the present sequence decodes to
XX the protein shown in ADA10874 not the one described in the specification
XX as being the nAChR alpha 7 subunit ADA10865.
XX
SQ Sequence 1876 BP; 369 A; 553 C; 531 G; 423 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,02e-124 Length: 1876
Score: 1258.50 Matches: 250
Percent Similarity: 63.23% Conservative: 75
Best Local Similarity: 48.64% Mismatches: 138
Query Match: 47.67% Indels: 51
DB: 8 Gaps: 7

US-09-303-232-6 (1-501) x ADA10864 (1-1876)
QY 8 LeuAlaLeuLeuAla----LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26
Db 100 CTGGCGCTGGCGCGCTGCTCTCGACGTGTCCTGCAAGCGAGTTCACAGGAAGCTT 159
QY 27 LeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSerGlu 46
Db 160 TACAAGGAGGTGGTCAAGAACTACAATCCCTTGGAGAGGCCGCTGGCAATGACTCGCA 219
QY 47 ProLeuGluValArgPheGlyLeuThrLeuGlnIleIleAspValAspGluLysAsn 66
Db 220 CCATCTACCGCTCTACTTCTCCCTGAGCTCTCTGACATCATGGACGTGATGAGAAGAC 279
QY 67 GlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArgTrp 86
Db 280 CAAGTTTAAACCAACATTTGGCTGCAATGTCTTGGACAGATCTATTTACAGTGG 339
QY 87 AsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThrProAsnLysLeuTrpLys 106
Db 340 AATGTGTCAAGATATCCAGGGGTGAAGACTGTTCTTCCAGATGGCCAGATTGGAAA 399
QY 107 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126
Db 400 CCAGACATCTCTCTATACAGTGTCTGATGAGCGCTTGGACGCCACATTCACACTAAC 459
QY 127 ValValValArgSerGlyGlySerCysLeuTyrValProGlyIlePheLysSerThr 146
Db 460 GTGTGTGGTAATCTTCTGGGCATTCGACGATTCCTCCCTCAGGCATATCAAGATTCC 519
QY 147 CysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMetLysPheGly 166
Db 520 TGCTACATCGATGATACGCTGGTTTCCCTTTGATGTGCAGCACTGCAAACTGAAGTTGG 579
QY 167 SerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyCys 186
Db 580 TCCTGGTCTTACGGAGGCTGCTCTTGGATCTGACATGTCAGAGGAG-----GCAGAT 630
QY 187 LeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGlyMetProGlyLysLysAsn 206
Db 631 ATCAGTGGCTATATCCCATGGAGATGGACCTAGTGGGAATCCCGCGCAAGAGGAGT 690
QY 207 ThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMetIle 226
Db 691 GAAAGGTTCTATGATGCTGCAAGAGGCCCTTACCCTGATGTCACCTTCACAGTACCATG 750
QY 227 ArgArgArgThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleSer 246

```

```

Db 751 CGCGCAGGAGCGCTCTACTATGGCTCAACCTGCTGATCCCTGTGTGCTCATCTCCGCC 810
QY 247 MetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuThrLeuGlyVal 266
Db 811 CTGCCCTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 870
QY 267 ThrLeuLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnVal 286
Db 871 ACAGTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 930
QY 287 SerAspAlaLeuProLeuLeuGlyThrTyPheAsnCysIleMetPheMetValAlaSer 306
Db 931 TCCGATTTCGTTACCATGATGACCCAGTACTTCCAGCAGCAGCATGATCATCTGTGGGCTC 990
QY 307 SerValValLeuThrValValValLeuAsnTyRhiShiArgThrAlaAspIleHisGlu 326
Db 991 TCGGTGTGTGTGATCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1050
QY 327 MetProGlnTrpIleLysSerValPheLeuGlnTrpLeuProTrpIleLeuArgMetSer 346
Db 1051 ATGCCAAGTGGACCAAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1110
QY 347 ArgProGlyLysIleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeu 366
Db 1111 AGGCGCGGAGGACAAAGTGGCGCGCGCTGCCAGCAGCAGCGCGCTGCAGCCTG 1170
QY 367 GluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAsp 386
Db 1171 GCCAGTGTGAGATGAGCGCGTGGCG-----GlyProGlyCys 407
QY 387 PheArgHisGlyProProProAsnSerThrAlaSerThrGlyAsnLeu----- 403
Db 1198 -----CGCGCGCC-----GCCAGCACGGGAACCTGCTGTATCATC 1233
QY 404 -----GlyProGlyCys 407
Db 1234 GGCTTCCGCGCGCTGGACGGGTGCTGTGTCCGACCCCGACTCTGGGGTAGTGT 1293
QY 408 SerIlePheArgThrAspPheArgSerPheValArgProSerThrMetGluAspVal 427
Db 1294 GGC-----CGCATGCGCTGTCTCCCGCACGACGATGAGCAGCTC 1332
QY 428 -----GlyGlyGlyLeuGlySerHisArgGluLeuHisLeuLeuArgGluLeu 445
Db 1333 CTGCACGGCGGGAACCCCGCGGGGACCCGACTTGGCCAGATCTCTGGAGAGTC 1392
QY 446 GlnPheIleThrAlaArgMetLysAlaAspGluAlaGluLeuLeuSerAspTrp 465
Db 1393 CGTACATTGCCAATCGCTTCCGCTGCCAGACGAGGCGGTCTGCAGCGAGTGG 1452
QY 466 LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle 485
Db 1453 AAGTTCGCGCGCTGTGTGTGGAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1512
QY 486 IleAlaThrValAlaValLeuLeuSerAlaProHisIleIle 499
Db 1513 ATCTGCACCATCGCATCTCTGATGTGGCTCCCACTCTGTG 1554

```

## RESULT 13

AAV44687

ID AAV44687 standard; cDNA; 1590 BP.

AC AAV44687;

XX 09-OCT-1998 (first entry)

XX V274T variant human alpha7 nAChR coding sequence.

XX Alpha7 nAChR; alpha7 nicotinic acetylcholine receptor subunit; cancer;  
 KW neurodegeneration; enzyme dysfunction; affective disorder; therapy;  
 KW immune dysfunction; diabetic neuropathy; Alzheimer's disease;  
 KW schizophrenia; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 9..1517

XX FT /\*tag= a

XX WO9828331-A2.

XX PN 02-JUL-1998.

XX PD 22-DEC-1997; 97WO-US023405.

XX PF 20-DEC-1996; 96US-00771737.

XX PR (ABBO ) ABBOTT LAB.

XX PA Briggs CA, Gopalakrishnan M, McKenna DG, Monteggia LM, Roch J;  
 XX PI Sullivan JP, Touma E;XX DR WPI; 1998-377593/32.  
 XX P-PSDB; AAW69216.

XX Nucleic acid encoding variant of human alpha 7 nicotinic acetylcholine  
 XX receptor sub-unit - used to identify modulators of the receptor,  
 XX potentially useful for treating neuro-degeneration, cancer etc.

XX Claim 14; Fig 2; 44pp; English.

XX This sequence encodes the V247T variant of human alpha7 nicotinic  
 XX acetylcholine receptor (nAChR) subunit of the invention. Cells containing  
 XX the DNA are used to express the protein and to identify modulators of  
 XX alpha7 nAChR activity or cytoprotective agents, e.g. antisense compounds  
 XX or antagonists that are potentially useful for treating  
 XX neurodegeneration, enzyme dysfunction, affective disorders and immune  
 XX dysfunction, such as cancer, post-herpetic neuralgia, diabetic  
 XX neuropathy, osteoarthritis, Alzheimer's or Parkinson's diseases, kuru,  
 XX psychosis and schizophrenia. Probes based on the DNA are used to detect  
 XX the DNA in usual hybridisation or amplification tests, while monoclonal  
 XX antibodies are used to detect the protein for diagnosis (in vitro or by  
 XX in situ immuno-fluorescent assay). Compared with wild-type alpha7 nAChR,  
 XX the protein has about 100-fold greater sensitivity to cholinergic  
 XX receptor agonists (nicotine or acetylcholine) and response to these  
 XX agonists decays more slowly, but the wild-type inward rectification is  
 XX retained

SQ Sequence 1590 BP; 322 A; 471 C; 449 G; 348 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	4,24e-124	Length:	1590
Score:	1254.50	Matches:	249
Percent Similarity:	63.04%	Conservative:	75
Best Local Similarity:	48.44%	Mismatches:	139
Query Match:	47.52%	Indels:	51
DB:	2	Gaps:	7

US-09-303-232-6 (1-501) x AAV44687 (1-1590)

QY 8 LeuAlaLeuLeuAla---LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26

Db 36 CTGGCACTGGCAGCATCTCTCTCTGTCCTGTCCTGCAAGCGCGAGTTCCAGAGGAGCTT 95

QY 27 LeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSerGlu 46

Db 96 TACAAGGAGCTGGTCAAGAACTCAATCCCTGGAGAGGCCCGTGGCAATGACTCGCAA 155

QY 47 ProLeuGluValArgPheGlyLeuThrLeuGlnIleIleAspValAspGluLysAsn 66

Db 156 CCACTACCGTCTACTTCTCTCCCTGAGGCTCTCTGAGCATATGACGCTGGATGAGAGAAC 215

QY 67 GlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArgTrp 86

Db 216 CAAGTTTTRACCAACCAATTTGGCTGCTTGGACAGATCACTATTACAGTGG 275



## Alignment Scores:

Pred. No.: 1.42e-123 Length: 2769  
 Score: 1253.00 Matches: 250  
 Percent Similarity: 64.65% Conservative: 81  
 Best Local Similarity: 48.83% Mismatches: 145  
 Query Match: 47.46% Indels: 36  
 DB: 2 Gaps: 8

US-09-303-232-6 (1-501) x AAT59196 (1-2769)

QY 2 AlaProMetLeuAlaLeuAlaLeuAlaLeuLeuProValSerGluGlnGlyPro 21  
 DB 83 GCGCTGATGCTGGCTGCTGGCGCGGGCTCGTGGCGAGTCCCTGCAGAGAG 142  
 QY 22 HisGluLysArgLeuLeuAsnAlaLeuAlaLeuLeuLeuLeuLeuLeuVal 41  
 DB 143 TTCAAAGGAAGCTGTCAAGGAGCTGTGAAGAACTACAACCCCTCTGGAACGACGTT 202  
 QY 42 AlaAsnGluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleAsp 61  
 DB 203 GCAATGACTCCCGCGCTGCTGCTATTTCACTCTCAGCCCTCATCGATCATGGAT 262  
 QY 62 ValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsnAsp 81  
 DB 263 GTGGATGAAGAATCAAGTATTACCAACAACTCTGGCTCAAAATGACTGGACAGAT 322  
 QY 82 TyrAsnLeuArgTrpAsnAspSerGluTrpGlyValLysAspLeuArgIleThrPro 101  
 DB 323 CATTTACTAGTGAATGTCTGAATACCTGGAGTGAAGACGTCCTTTTCCTGAT 382  
 QY 102 AsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly 121  
 DB 383 GGAAGTATTGGAAGCCAGATATCTCTATACAGTCTGTATGAAGATTTGATGCT 442  
 QY 122 ThrTyrGlnThrAsnValValArgSerGlyGlySerCysLeuTyrValProGly 141  
 DB 443 ACATTTACACTAATGTTTAGTCAATCTCGGACACTGCCAATATCGCCACACGCG 502  
 QY 142 IlePheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspGlnHisCys 161  
 DB 503 ATATTTAAAGCTCATGCTACATAGACGCTGTTGTTTCCATTTGATGTTTCAAGAGTGC 562  
 QY 162 AspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLysAsp 181  
 DB 563 AATCTGAAGTTGATCTTGACATATGGAGCTGGTCTTAGACTTTACAAATGCAAGAA 622  
 QY 182 GluAlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGlyMet 201  
 DB 623 -----CGAGATATATCGGCTATATTTCAAATGGAGAGTGGGATTTAGTAGNAAT 673  
 QY 202 ProGlyLysAsnThrIleThrTyrAlaCysCysProGluProTyrValAspValThr 221  
 DB 674 CTTGGGAAGAGAACTGAGAGCTTTATGAGTGTGTAAAGAACCATACCCAGATATCACA 733  
 QY 222 PheThrIleMetIleArgArgGlyThrLeuTyrTyrPhePheAsnLeuIleValProCys 241  
 DB 734 TTCACAGTAAACCATGAGACGACAGACTCTTACTACGGGCTCAACCTTCTATTTCCTGT 793  
 QY 242 ValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLys 261  
 DB 794 GTACTGATATCAGCACTTGCTATTATAGTCTTTCTGCTCCAGCAGACTCAGAGANAG 853  
 QY 262 LeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGlu 281  
 DB 854 ATCTCAGTAGTATAACAGATTTATTGCTCTCACCGTCTCATGTTTACTCGTGGCTGAA 913  
 QY 282 ThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMet 301  
 DB 914 ATTATGCCAGCAACATCTGATTCTGTGCCCTTAATGCTCTCAGTATTTTGCAGCACCATG 973  
 QY 302 PheMetValAlaSerValValLeuThrValValLeuAsnTyrHisHisArgThr 321

DB 974 ATTATTGTTGGCCTCTCTGTTGTTGTGTCACGTGTATGCTTACAAATCCATCAGAT 1033  
 QY 322 AlaAspIleHisGluMetProGlnTrpIlePheLysSerValPheLeuGlnTrpLeuProTrp 341  
 DB 1034 CCAGATGGGGAAATGCTAAATGCAAGAGTCACTCTTCTGAATTTGGTGTGCTGG 1093  
 QY 342 IleLeuArgMetSerArgProGlyLysIleThrArgLysThrIleMetMetAsnThr 361  
 DB 1094 TTTCTGAGGATGAAGACAGCAGGGAAGATAAAGTGGTCCCGCTCTCAACATAAAG 1153  
 QY 362 ArgMetArgGluLeuGluLeuLysGluArgSerSerLysSer-----LeuLeu 377  
 DB 1154 CGCGATGCGACCTGTCAAGCATGGAGATGAACACTGTGAGTGTGTCAGCATGAGTAAT 1213  
 QY 378 AlaAsnValLeuAspIleAspAspPheArg-----HisGlyProPro 392  
 DB 1214 GGGAAACATGCTGTATATT-----GGGTTTCGAGGGCTGGATGGGTTCATCTGCACACCC 1267  
 QY 393 ProProAsnSerThrAlaSerThrGlyAsnLeuGlyProGlyCysSerIlePheArgThr 412  
 DB 1268 ACCACTGATTCAAGGGGTGATCTCTGGAGGATG-----ACCTGTTTCA----- 1309  
 QY 413 AspPheArgArgSerPheValArgProSerThrMetGluAspValGlyGlyLeuGly 432  
 DB 1310 -----CCACAGAGAGAAATCTT-----CTGCAC 1336  
 QY 433 SerHisHis-----ArgGluLeuHisLeuIleLeuArgGluLeuGlnPhe 447  
 DB 1337 AGTGGCCACCCCTCTGAAGCGCACCCAGATTGCTTAAGATCTTGAAGAGGTCAGATAC 1396  
 QY 448 IleThrAlaArgMetLysLysAlaAspGluAlaGluLeuLeuSerAspTrpLysPhe 467  
 DB 1397 ATTGCAAAACAGGTTCAAGACACAGGATGAAGAAAGACCATTTGCACGAATGGAAGTTT 1456  
 QY 468 AlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIleAla 487  
 DB 1457 GCAGCCTCTGTAGTAGATCGCTCTGCTTGTATGTCATTTTCGGTCTTCCATCATTTGT 1516  
 QY 488 ThrValAlaValLeuLeuSerAlaProHisIleIle 499  
 DB 1517 ACAATTGGCATCTTAATGTGCAGCACCAACTTTGTA 1552

## RESULT 15

AAC90385  
 ID AAC90385 standard; cDNA; 1509 BP.

XX AAC90385;  
 AC AAC90385;  
 DT 14-MAR-2001 (first entry)  
 XX Mutant human alpha7 ligand gated ion channel coding sequence #1.  
 DE Human; alpha7 nicotinic acetylcholine gated ion channel; mutant;  
 KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance; ss.  
 XX Homo sapiens.

XX WO200073431-A2.

XX 07-DEC-2000.

XX 25-MAY-2000; 2000WO-US011862.

XX 27-MAY-1999; 99US-0136174P.

XX (PHAA ) PHARMACIA & UPJOHN CO.

XX Groppi VB, Wolfe ML, Berkenpas MB;

XX WPI; 2001-061524/07.

XX P-PSDB; AAB50015.

XX Special cell culture medium for treating cells and for inducing mammalian

PT cell lines to conduct calcium ions, comprising specified concentrations  
 PR of ions of sodium, calcium and potassium at specified pH.  
 XX  
 PS Claim 50; Page 69; 77pp; English.  
 XX

CC The present sequence is the coding sequence for a mutant human alpha7  
 CC nicotinic acetylcholine gated ion channel. The human alpha7 ion channel  
 CC was used in the construction of an alpha7/5-hydroxytryptamine (5-HT3)  
 CC chimeric ligand gated ion channel (see AAC90382 and AAB50014). The  
 CC alpha7/5-HT3 chimeric ion channel can be expressed by recombinant cells  
 CC in the present invention, resulting in preferential calcium ion  
 CC conductance by the cells. The protein encoded by this sequence has the  
 CC wild-type threonine residue at position 230 substituted by a proline  
 CC residue

XX Sequence 1509 BP; 296 A; 452 C; 429 G; 332 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.: 6.42e-124 Length: 1509  
 Score: 1252.50 Matches: 249  
 Percent Similarity: 63.04% Conservative: 75  
 Best Local Similarity: 48.44% Mismatches: 139  
 Query Match: 47.44% Indels: 51  
 DB: 4 Gaps: 7

US-09-303-232-6 (1-501) x AAC90385 (1-1509)

QY 8 LeuAlaLeuLeuAla--LeuLeuProValSerGluGlnGlyProHisGluLysArgIeu 26  
 DB 28 CTGGCGCTGGCGGTGCTCTCTGACGTGCTCTGCAAGCGAGTTCAGAGGAAGCTT 87  
 QY 27 LeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSerGlu 46  
 DB 88 TACAAGGAGCTGGTCAAGACTACATCCCTTGGAGAGCCGCTGGCCATGACTCGAA 147  
 QY 47 ProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleLeuAspValAspGluLysAsn 66  
 DB 148 CCATCTACCGTCTACTTCTCTGAGCTCTCTGAGATCATGAGCGTGGATGAGAAGAA 207  
 QY 67 GlnLeuLeuThrAsnIleThrLeuSerLeuGluThrAsnAspTyrAsnLeuArgTrp 86  
 DB 208 CAAGTTTAAACCAACATTTGGCTGCAATGCTTGGACAGATCATATTTACAGTGG 267  
 QY 87 AsnAspSerGluTyrGlyValLysAspLeuArgIleThrProAsnLysLeuTrpLys 106  
 DB 268 AATGTGTCAAGATATCCAGGGGTGAAGACTGTCTGTTCCAGATGCCAGATTTGGAA 327  
 QY 107 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126  
 DB 328 CCAGACATTTCTCTATAACAGTGTCTGATGAGCGCTTTGACGCCACATTCACACTAAC 387  
 QY 127 ValValValArgSerGlySerCysLeuTyrValProGlyIlePheLysSerThr 146  
 DB 388 GTGTGTGTGAATTTCTTGGGCATTCGACGATTCCTCCAGCAATATTCAGAGTTCC 447  
 QY 147 CysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMetLysPheGly 166  
 DB 448 TGCTACATCGATGACGTGTTTCCCTTTGATGTCAGCACTGCACAACTGAAGTTTGGG 507  
 QY 167 SerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyLys 186  
 DB 508 TCCCTGGTCTTACGAGGCTGCTCTGATTCGAGATGTCAGAG-----GCAGAT 558  
 QY 187 LeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGlyMetProGlyLysLysAsn 206  
 DB 559 ATCAGTGGCTATATCCCAATGGAGATGGACCTAGTGGGAATCCCGCGCAAGAGGAGT 618  
 QY 207 ThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMetIle 226  
 DB 619 GAAAGGTTCTATGATGTGCAAAAGAGCCCTTACCCCGATGTCACTTCACAGTACCATG 678  
 QY 227 ArgArgArgThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleSerSer 246

DB 679 CGCGCAGGCGCGCTTACTATGCGCTCAACCTGTGTATCCCTGTGTCTCATCTCCGCC 738  
 QY 247 MetAlaLeuLeuGlyPheThrLeuProAspSerGlyGluLysLeuThrLeuGlyVal 266  
 DB 739 CTCGCCCTGTGGTTCCTCTCTGAGATTCCGGGGAAGATTTCCTCGGGGATA 798  
 QY 267 ThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnVal 286  
 DB 799 ACAGTCTTACTCTCTTACCGTCTTCATGCTGCTGCTGGCTGAGATCATGCCCGCAACA 858  
 QY 287 SerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSer 306  
 DB 859 TCCGATTTCGTTACCATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 918  
 QY 307 SerValValLeuThrValValValLeuAsnTyrHisArgThrAlaAspIleHisGlu 326  
 DB 919 TCGGTGTGTGGTGGACGGTGTCTGCTGAGTATGATGATGATGATGATGATGATGATGATG 978  
 QY 327 MetProGlnTrpIleLysSerValPheLeuGlnTrpLeuProTrpIleLeuArgMetSer 346  
 DB 979 ATGCCCAAGTGGACGAGTATCTCTTGAATGCTGCGGTGGTCTCTCGAATGAAG 1038  
 QY 347 ArgProGlyLysLysIleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeu 366  
 DB 1039 AGCCCGGGGAGGACAAAGGTGCGCCCGCTGCCAGCACAGCAGCGCGCTGCAGCGCTG 1098  
 QY 367 GluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAsp 386  
 DB 1099 GCCAGTGTGGAGTGGACGGCGGTGGCG----- 1125  
 QY 387 PheArgHisGlyProProProProAsnSerThrAlaSerThrGlyAsnLeu----- 403  
 DB 1126 -----CCGCGCGCC-----GCCAGCAACGGGAACCTCTGTGTACATC 1161  
 QY 404 -----GlyProGlyCys 407  
 DB 1162 GGCTTCGCGCGCTGGACGGCGTGCATGTTCCCGACCCCGCTCTGGGTGTGTGTGT 1221  
 QY 408 SerIlePheArgThrAspPheArgSerPheValArgProSerThrMetGluAspVal 427  
 DB 1222 GGC-----CGCATGGCTGTCTCCCGACCGCAGCATGACACCTC 1260  
 QY 428 -----GlyGlyGlyLeuGlySerHisArgGluLeuHisLeuLeuArgGluLeu 445  
 DB 1261 CTGCACGGCGGCAACCCCGAGGGGACCCCGACTTGGCCAAAGATCTCTGGAGAGTGC 1320  
 QY 446 GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuIleSerAspTrp 465  
 DB 1321 CGCTACATTCCTGCTTCCGCTGCCAGCAGAAAGCGAGCGGTCTGCAGCGAGTGG 1380  
 QY 466 LysPheAlaAlaMetValValAspPheCysLeuPheValPheThrLeuPheThrIle 485  
 DB 1381 AAGTTCCGCGCTGTGTGTGGACCGCTGTGCTCTATGGCTTCTCGGTCTTCACCATC 1440  
 QY 486 IleAlaThrValAlaValLeuLeuSerAlaProHisIleLeile 499  
 DB 1441 ATCTGCACCATCGGCATCTCTGATGCTCGCTCCCAACTCGTG 1482

Search completed: May 8, 2004, 19:30:45  
 Job time : 554.234 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:31:03 ; Search time 12.1919 Seconds  
(without alignments)  
3952.801 Million cell updates/sec

Title: US-09-303-232-6  
Perfect score: 2640  
Sequence: 1 MAPMLAALALLPVSSEQ.....LFTIATVALLSAPHIVQ 501

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: Piri:\*  
2: Piri2:\*  
3: Piri3:\*  
4: Piri4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1260	47.7	502	1 ACHUA7	nicotinic acetylch
2	1255.5	47.6	502	2 A57175	nicotinic acetylch
3	1253	47.5	502	2 JN0113	nicotinic acetylch
4	1252.5	47.4	502	2 G02259	alpha 7 neuronal n
5	1246.5	47.2	502	2 T01378	nicotinic receptor
6	1237.5	46.9	498	2 S68588	nicotinic acetylch
7	1226.5	46.5	511	2 JH0173	alpha-bungarotoxin
8	1132	42.9	461	2 T25671	hypothetical prote
9	1083.5	41.0	542	2 T19862	hypothetical prote
10	1081	40.9	560	2 T19622	hypothetical prote
11	989	37.5	503	2 A53956	nicotinic acetylch
12	982.5	37.2	502	2 A37040	nicotinic acetylch
13	970.5	36.8	495	2 S60589	acetylcholine rece
14	970.5	36.8	499	2 A24572	nicotinic acetylch
15	964.5	36.5	567	1 ACFPA1	nicotinic acetylch
16	944	35.8	557	2 S12359	nicotinic acetylch
17	943	35.7	576	1 ACFPA2	nicotinic acetylch
18	929.5	35.2	494	2 T09289	nicotinic acetylch
19	924.5	35.0	528	1 AACH2N	nicotinic acetylch
20	919.5	34.8	512	2 B37014	nicotinic acetylch
21	919	34.8	511	2 A40110	nicotinic acetylch
22	910	34.5	500	2 S12899	nicotinic acetylch
23	909.5	34.5	495	2 B35721	nicotinic acetylch
24	906.5	34.3	517	2 A30992	probable nicotinic
25	906	34.3	457	1 ACBOA1	nicotinic acetylch
26	905	34.3	457	1 ACHUA1	nicotinic acetylch
27	901	34.1	521	1 ACFENN	nicotinic acetylch
28	900	34.1	503	2 JH0174	nicotinic acetylch
29	897.5	34.0	627	2 JC4021	nicotinic acetylch

## ALIGNMENTS

RESULT 1  
ACHUA7

nicotinic acetylcholine receptor alpha-7 chain precursor, neuronal - human  
N;Alternate names: cholinergic nicotinate receptor alpha-7 chain  
C;Species: Homo sapiens (man)  
C;Date: 12-Aug-1996 #sequence revision 31-Jan-1997 #text\_change 22-Jun-1999  
C;Accession: I37185; A54194; S60309  
R;Peng, X.; Katz, M.; Gerzanich, V.; Anand, R.; Lindstrom, J.  
Mol. Pharmacol. 45, 546-554, 1994  
A;Title: Human alpha 7 acetylcholine receptor: cloning of the alpha 7 subunit from the s  
omers expressed in Xenopus oocytes.  
A;Reference number: I37185; MUID:94195283; PMID:8145738  
A;Accession: I37185  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-502 <PEN>  
A;Cross-references: EMBL:X70297; NID:g496606; PIDN:CAA49778.1; PID:g496607  
A;Experimental source: brain neuroblastoma cell line SHSY-5Y  
R;Chini, B.; Raimond, E.; Elgoyhen, A.B.; Moralli, D.; Balzaretti, M.; Heinemann, S.  
Genomics 19, 379-381, 1994  
A;Title: Molecular cloning and chromosomal localization of the human alpha-7-nicotinic  
A;Reference number: A54194; MUID:94245214; PMID:8188270  
A;Accession: A54194  
A;Molecule type: mRNA

30	896	33.9	456	1	ACCHAN	nicotinic acetylch
31	895	33.9	502	2	S10505	nicotinic acetylch
32	891	33.8	622	1	ACCH4N	nicotinic acetylch
33	890	33.7	457	2	S13872	nicotinic acetylch
34	890	33.7	470	2	A39218	nicotinic acetylch
35	888	33.6	625	2	A26456	nicotinic acetylch
36	887	33.6	498	2	G02421	nicotinic acetylch
37	885.5	33.5	459	2	S14703	nicotinic acetylch
38	885	33.5	457	2	A24383	nicotinic acetylch
39	883.5	33.5	445	2	I49458	nicotinic acetylch
40	881	33.4	491	1	ACHHNN	acetylcholine rece
41	879.5	33.3	461	2	I50548	acetylcholine rece
42	876.5	33.2	461	1	ACRFA1	nicotinic acetylch
43	875.5	33.2	457	2	A28529	nicotinic acetylch
44	869.5	32.9	457	2	S08162	nicotinic acetylch
45	862	32.7	559	2	E89134	protein F2566.4 [i

Query Match 47.7%; Score 1260; DB 1; Length 502;  
Best Local Similarity 48.5%; Pred. No. 2.9e-99;  
Matches 247; Conservative 75; Mismatches 137; Indels 50; Gaps 6;



A;Residues: 24-25,'ET',28-41,'X',43-45,'X',47 <CON>  
C;Comment: This acetylcholine receptor is blocked by alpha-bungarotoxin and is localized  
C;Genetics:  
A;Introns: 19/1; 65/3; 80/3; 117/2; 144/1; 200/1; 265/1; 294/1; 330/3  
C;Superfamily: acetylcholine receptor  
C;Keywords: brain; glycoprotein; ion channel; neurotransmitter receptor; phosphoprotein;  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-502/Product: nicotinic acetylcholine receptor alpha-7 chain, neuronal #status predi  
F;231-254/Domain: transmembrane #status predicted <TR1>  
F;262-280/Domain: transmembrane #status predicted <TR2>  
F;296-317/Domain: transmembrane #status predicted <TR3>  
F;470-488/Domain: transmembrane #status predicted <TR4>  
F;46,90,133/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;365,367,413,427,465/Binding site: phosphate (Ser) (covalent) #status predicted  
F;415/Binding site: phosphate (Thr) (covalent) #status predicted  
F;442/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 47.5%; Score 1253; DB 2; Length 502;  
Best Local Similarity 48.8%; Pred. No. 1.1e-98;  
Matches 250; Conservative 81; Mismatches 145; Indels 36; Gaps 8;

QY 2 APMALALALLPVSEOGPHEKLLNALLANYNTLPRVANSEPLEVRFGTLTQQIID 61  
Db 5 ALMLLAAAGLVRESLQGEFQKLYKELKNYNPLRPVANDSQPLTVYFTLSLMQIMD 64

QY 62 VDEKNQLITNIWLSLEWYNDYLRWDSYGGVKDLRITPNKLMKPDVLMYNSADEGDFG 121  
Db 65 VDEKNQVLTNIWLMQWTDHYLQWNVSEYGVKVRFPDGLTWKPDILLYNSADERFDA 124

QY 122 TYQTNVVRSGGCLVPPGIFKSTCKMDIAWFPDDOHCMDKFGSWTYDGNOLDIVLKD 181  
Db 125 TFEHNVVNSGHCQVLPFGIFKSSCYIDVRWFPDVQKCNLFKFGSWTYGWSLDLQME 184

QY 182 EAGGDLSDFTNGEWLIGMPGKNTITYACCPYVDVFTTMRRLTYFFNLIVPC 241  
Db 185 --ADISGYISNGEWDLVGPGKRTSEYECCKEPYDITFTVTRRLTYGLNLLIPC 241

QY 242 VLISSVALLGFTLPPDSGEKLTGLVTLLSLTVFLNLVAETLPQVSDAIPLLGYFNCIM 301  
Db 242 VLSALALLVFLPADSGEKISLIGITVLLSLTVFMLVAEIMPATSDSVPLIAQYFAS 301

QY 302 FMVASSVLTIVVLYNHRHTADIEHMPQWIKSVFLOLWPLILMSRPGKKITRKTMMNT 361  
Db 302 IIVGLSVVTVIVLQHHHPDGGKMPKTRVILLNCWFLMKRPGEDKVRPAQCHKQ 361

QY 362 RMRELKERSKS---LLANVLIDDDPR---HGPPPPNSTASTGNLPGCSIFRT 412  
Db 362 RRCSSLMEMNTVSGQCSNGNMLYI--GPRGLDGVHCTPTTDSGVICGRM--TCS--- 413

QY 413 DFRSFRVPSMTMEDVGGGLGSHH-----RELHLILRELOITARMKKADEAELISDWKF 467  
Db 414 -----PTEENL-----LHSGHPSEGDPLAKILEEVRYIANFRQDSEEAICNEKF 462

QY 468 AAMVDRFCIFVETFTIATVAVLLSAPHII 499  
Db 463 AASVDELCLMAFSVFTICTIGILMSAPNFV 494

RESULT 4  
G02259  
alpha 7 neuronal nicotinic acetylcholine receptor - human  
C;Species: Homo sapiens (man)  
C;Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 20-Aug-1999  
C;Accession: G02259  
R;Leonard, S.  
submitted to the EMBL Data Library, November 1995  
A;Accession: G02259  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-502 <LEO>  
A;Cross-references: EMBL:U40583; NID:g1125076; PIDN:AAA83561.1; PID:g1125077  
C;Superfamily: acetylcholine receptor

Query Match 47.4%; Score 1252.5; DB 2; Length 502;  
Best Local Similarity 48.4%; Pred. No. 1.3e-98;  
Matches 249; Conservative 74; Mismatches 140; Indels 51; Gaps 7;

QY 8 LALLA-LLPVSEOGPHEKLLNALLANYNTLPRVANSEPLEVRFGTLTQQIIDVDEKN 66  
Db 10 LALAALLHVSUGEFQKLYKELKNYNPLRPVANDSQPLTVYFSLNLQIMDVDEKN 69

QY 67 QLLITNIWLSLEWYNDYLRWDSYGGVKDLRITPNKLMKPDVLMYNSADEGDFGTQTN 126  
Db 70 QVLTNNIWLQMSWTDHYLQWNVSEYGVKVRFPDGLTWKPDILLYNSADERDAFHTN 129

QY 127 VVVRSGGCLVPPGIFKSTCKMDIAWFPDDOHCMDKFGSWTYDGNOLDIVLKDAGGD 186  
Db 130 VLVPNGHCQVLPFGIFKSSCYIDVRWFPDVQKLFKFGSWYGGWSLDLQMQE---AD 186

QY 187 LSPFITNGEWLIGMPGKNTITYACCPYVDVFTTMRRLTYFFNLIVPCVLIS 246  
Db 187 ISGYIENGWDLVIGIFKRSERYECKEPYDITFTVTRRLTYGLNLLIPCULISA 246

QY 247 MALLGFTLPPDSGEKLTGLVTLLSLTVFLNLVAETLPQVSDAIPLLGYFNCIMFVAS 306  
Db 247 LALLVFLPADSGEKISLIGITVLLSLTVFMLVAEIMPATSDSVPLIAQYFAS 306

QY 307 SVVLTIVVLYNHRHTADIEHMPQWIKSVFLOLWPLILMSRPGKKITRKTMMNTMREL 366  
Db 307 SVVTVIVLQYHHHPDGGKMPKTRVILLNCWFLMKRPGEDKVRPAQCHKQRCSL 366

QY 367 ELKERSKLLANVLIDDDFRHGPPEPNSTAGNL-----GPGC 407  
Db 367 ASVEMSAVA-----PPP---ASGNLLYIGRGLDGVHCVPTPDSGVVC 407

QY 408 SIFRTDPRFSVRPSTMEDV--GGGLGSHHRELHLILRELOITARMKKADEAELISDW 465  
Db 408 G-----RMACSPTHDEHLHGGQPEGDPDLAKILEEVRYIANFRQDSEEA 460

QY 466 KFAAMVDRFCLVFTIATVAVLLSAPHII 499  
Db 461 KFAACVVDRLCLMAFSVFTICTIGILMSAPNFV 494

RESULT 5  
T01378  
nicotinic receptor alpha 7 chain - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 21-Jul-2000  
C;Accession: T01378  
R;Seguela, P.; Wadiche, J.; Dineley-Miller, K.; Dani, J.A.; Patrick, J.W.  
J. Neurosci. 13, 596-604, 1993  
A;Title: Molecular cloning, functional properties, and distribution of rat brain alpha  
A;Reference number: Z14310; MUID:93147931; PMID:7678957  
A;Accession: T01378  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-502 <SEG>  
A;Cross-references: EMBL:S53987; NID:g264770; PIDN:AAB25224.2; PID:g5705903  
A;Experimental source: brain  
C;Superfamily: acetylcholine receptor

Query Match 47.2%; Score 1246.5; DB 2; Length 502;  
Best Local Similarity 49.4%; Pred. No. 4.1e-98;  
Matches 248; Conservative 78; Mismatches 149; Indels 27; Gaps 7;

QY 8 LAL-LLPVSEOGPHEKLLNALLANYNTLPRVANSEPLEVRFGTLTQQIIDVDEKN 66  
Db 10 LALAALLHVSUGEFQKLYKELKNYNPLRPVANDSQPLTVYFSLNLQIMDVDEKN 69

QY 67 QLLITNIWLSLEWYNDYLRWDSYGGVKDLRITPNKLMKPDVLMYNSADEGDFGTQTN 126  
Db 70 QVLTNNIWLQMSWTDHYLQWNVSEYGVKVRFPDGLTWKPDILLYNSADERDAFHTN 129

QY 127 VVVRSGGCLVPPGIFKSTCKMDIAWFPDDOHCMDKFGSWTYDGNOLDIVLKDAGGD 186

Db 130 VLVNASGHCQYLPFGIFKSSCYIDVRWFFPDVQCKLKFGSWSYGSGLDLQMQE---AD 186  
Qy 187 LSDPITNGEWYLGMPGKKNITVACCEPVDVFTIMIRRTLYFFNLIIVPCVLIS 246  
Db 187 ISSYIPNGEWMGIPGRNKFYECCKEPPDVTYVTMERRTLYYGLNLIIPCVLISA 246  
Qy 247 MALLGFTLPPDSGKLTGLVTLLSLVFLNLVAETLPQVSDAIPLLGTYENCIMFMVAS 306  
Db 247 LALLVFLPADSGKISLGITVLSLVFLNLVAETLPQVSDAIPLLGTYENCIMFMVAS 306  
Qy 307 SVLITVVVLYNHHRTADITHMPQWIKSVFLQWLPWILMSRPGKKITRKTIMMTRMREL 366  
Db 307 SWVTIVLVYHHDPDGGKPKWTRIIILLNCAWFLRMKRPGEKVPAQCHKPRCSL 366  
Qy 367 ELKERS-----KSLIANVLIDDDFR-----HGPPPNSTASTGNLPGCCSIFRTDFRS 417  
Db 367 ASVELSAGAGPTNGNLLYI--GFRLEGHMCAPTDPGVCVGRLL--ACSPTHDEHLMH 422  
Qy 418 FVRPSTMEDVGGGLGSHHRELHLILRELQFITARMKKADEBAELISDWKFAAMVVDPRCL 477  
Db 423 GAHPSDGP-----DLAKILEEVRYIANRNCODESEVICSEWKFAACVVDPLCL 472  
Qy 478 FVFTLFTIIATVALLSAPHII 499  
Db 473 MAFSVFTIICIGILMSAPNFV 494

## RESULT 6

S68388  
nicotinic acetylcholine receptor alpha-1 chain precursor (clone Ce21) - Caenorhabditis e  
C:Species: Caenorhabditis elegans  
C:Date: 06-Dec-1996 #sequence\_revision 07-Feb-1997 #text\_change 20-Aug-1999  
C:Accession: S68588; S57496  
R:Ballivet, M.; Alliod, C.; Bertrand, S.; Bertrand, D.  
J. Mol. Biol. 258, 261-269, 1996  
A:Title: Nicotinic acetylcholine receptors in the nematode Caenorhabditis elegans.  
A:Reference number: S68587; MUID:96196478; PMID:8627624  
A:Accession: S68588  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-498 <BAL>

A:Cross-references: EMBL:X83887; NID:9872087; PIDN:CAA58764.1; PID:g872088

C:Superfamily: acetylcholine receptor

C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membrane;

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-498/Product: nicotinic acetylcholine receptor alpha-1 chain #status predicted <MAT>

Query Match 46.9%; Score 1237.5; DB 2; Length 498;  
Best Local Similarity 47.5%; Pred. No. 2.4e-97;  
Matches 245; Conservative 80; Mismatches 140; Indels 51; Gaps 9;

Qy 2 APLAALLALLPVSEGGPHEKRLNALLANNYTLERPVAENSEPLEVRFGLTLQIID 61  
Db 15 APLGSL-----QERLYEDLMRYNNLRPNVANHSEPVTVHLKVALQIID 61  
Qy 62 VDEKNQLITNIWLSLEWNDYLNWDSYGVKDLRITNKLWKPDLVMTNSADEGFDG 121  
Db 62 VDEKNQVVVNAWLDYTWNDYLVMDKAEYGNITDVRFPAGKIWKPDVLLNYSVDTFDS 121  
Qy 122 TYQTNVVVRSGGCLYVPPGIFKSTCKMDIAWFFPDQCHDKMGKSWTYDGNQLDLVKD 181  
Db 122 TYQTNMIVYSGLVVHWPVPGIFKLSCKID:QWFFPDQCKFFKFGSWTYDGYKLD--LQP 179  
Qy 182 EAGG-DLSDFTNGEWYLGMPGKKNITVACCEPVDVFTIMIRRTLYFFNLIIVP 240  
Db 180 ATGCFD:SEYISNGEWALPLTTVERNEKFFDCCPEPVDVHYLHMRRRTIYGFNLMP 239  
Qy 241 CVLISMAILGFTLPPDSGKLTGLVTLLSLVFLNLVAETLPQVSDAIPLLGTYENCI 300  
Db 240 CILATLMTLLGFTLPPDAGEKITIQTIVLLSICFFLSIVSEMSPTSEAVPLLIIFTCC 299  
Qy 301 MFMVASSVLTVVVLYNHHRTADITHMPQWIKSVFLQWLPWILMSRPGKKITRKTII-M 359

Db 300 MIVVTASTVFTVYVNLHLHYRTPETHDMPWRNLLLYWILRMKRGHNLTYASLPSL 359  
Qy 359 MNTMRRELEUKERSKSLANVLD-----IDDFRHGPPPNSTASTGNLPGPCS 408  
Db 360 PSTK-----PNRHSSEILIRIKONEHSLSRANSFDADCRNLQVIMTQSVNGLTSLG-S 412  
Qy 409 IFRDTFRSFRVPRSTMEDVGGGLGSHHRE-----LHLILRELQFITARMKKADEBAELIS 463  
Db 413 I-----PSTMISSNGTTTVDVSOQATLLILHRIYHELKIVTRKMIEGDKBEQACN 461  
Qy 464 DWKEFAAMVVDPRCLFVFTLFTIIATVALLSAPHII 499  
Db 462 NWKFAAMVVDRLCLYVFTIIVTIGIFWNSAPVIV 497

## RESULT 7

JH0173

alpha-bungarotoxin-binding protein alpha-2 chain precursor - chicken

C:Species: Gallus gallus (chicken)

C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 20-Aug-1999

C:Accession: JH0173

R:Schoepfer, R.; Conroy, W.G.; Whiting, P.; Gore, M.; Lindstrom, J.

Neuron 5, 35-48, 1990

A:Title: Brain alpha-bungarotoxin binding protein cDNAs and WABs reveal subtypes of thi

A:Reference number: JH0172; MUID:90315158; PMID:2369519

A:Accession: JH0173

A:Molecule type: mRNA

A:Residues: 1-511 <SCH>

A:Cross-references: GB:X52296; NID:G63081; PIDN:CAA36544.1; PID:g63082

A:Experimental source: brain

A:Note: this sequence is similar to acetylcholine receptor alpha chains

C:Comment: Alpha-bungarotoxin binding proteins are localized to extrasynaptic pseudoden

C:Superfamily: acetylcholine receptor

C:Keywords: glycoprotein; transmembrane protein

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-510/Product: alpha-bungarotoxin binding protein alpha-2 chain #status predicted <AL

F:239-262/Domain: transmembrane #status predicted <TM1>

F:270-288/Domain: transmembrane #status predicted <TM2>

F:304-323/Domain: transmembrane #status predicted <TM3>

F:479-496/Domain: transmembrane #status predicted <TM4>

F:54/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 46.5%; Score 1226.5; DB 2; Length 511;

Best Local Similarity 46.6%; Pred. No. 2.1e-96;

Matches 246; Conservative 75; Mismatches 134; Indels 73; Gaps 9;

Qy 5 LAALALLALLPVSEGGPHEKRLNALLANNYTLERPVAENSEPLEVRFGLTLQIIDVDE 64

Db 16 LWASLFLSFFKVSQGGESQRRLYRDLRLNRYNLERPVWDSQPIVVELQLSLQIIDVDE 75

Qy 65 KNQLITNIWLSLEWNDYLNWDSYGVKDLRITNKLWKPDLVMTNSADEGFDGTYQ 124

Db 76 KNQVLIITNAWLQMTVWDYILSWDQYEQYQVQNLRFPSDQIWPVDILLYNSADERFDATFH 135

Qy 125 TNVVRSGGCLYVPPGIFKSTCKMDIAWFFPDQCHDKMGKSWTYDGNQLDLVKDEAG 184

Db 136 TNVLVNYSGCQYIPGILKSTCYIDVRWFFPDVQCKDLKFGSWTHSGWLLDQMLE--- 192

Qy 185 GDLSDFITNGEWYLGMPGKKNITVACCEPVDVFTIMIRRTLYFFNLIIVPCVLI 244

Db 193 ADISNYISNGEWDLVGVGPKRNLRYECCKEPPDVTYVTMERRTLYYGLNLIIPCVLI 252

Qy 245 SSMAILGFTLPPDSGKLTGLVTLLSLVFLNLVAETLPQVSDAIPLLGTYENCFMV 304

Db 253 SGIALVFLPADSGEKISLGITVLLSLVFLNLVAETLPQVSDAIPLLGTYENCFMV 312

Qy 305 ASSVLTVVVLYNHHRTADITHMPQWIKSVFLQWLPWILMSRPGKKITRKTIMMTRMR 364

Db 313 GLSVVTVLVLFQHHDPQAGKMPWRVVRILLNCAWFLRMKKEGNI-----X 361

Qy 365 ELELK-----ERSKSLANVLIDDDFRHGPPPNSTASTGNL-----GPGGSIF 410



R:Kershaw, J.  
submitted to the EMBL Data Library, April 1997  
A:Reference number: Z19153  
A:Accession: T19622  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-560 <WIL>  
A:Cross-references: EMBL:Z93778; PID:CAB07843.1; GSPDB:GN00019; CESP:C31H5.3  
A:Experimental source: clone C31H5  
C:Genetics:  
A:Gene: CESP:C31H5.3  
A:Map position: 1  
A:Introns: 24/1; 70/3; 139/2; 270/1; 299/1; 336/3; 372/2; 456/3  
C:Superfamily: acetylcholine receptor

Query Match.	40.9%;	Score 1081;	DB 2;	Length 560;
Best local Similarity	39.4%;	Pred. No. 6e-84;		
Matches 218;	Conservative 88;	Mismatches 155;	Indels 92;	Gaps 8;
Qy	20	GPHEKRLNLNALLANYNTLERPVANESRPLEVRFGLTLQOIIDVDEKQOLLITNWLISLEW	79	
Db	28	GDHERRLYAKLAENYNKLAPVRNESRAVVLLGMDYQOIILIDIEKQIMNSNVNLRNSW	87	
Qy	80	NDYNLRWNSYGGVGVKOLRIITPNKLPDVLMYNSADEGFGDTYTNNVRSGGSCSLYVP	139	
Db	88	TDHYLTWDPSEFGNIKEVRLPINNIWKPDVLLYNSVQDPDSTWPVNAVVLTYGTNVTWIP	147	
Qy	140	PGIFKSPCKWDIAWFPPDDOCHDMKFGSWTDGNQLDLVLKDEAGGDLSDPITNGEWLI	199	
Db	148	PAIRSRSCAIDIAVFPFDTOHCTWKFGSWTYSGFFTLI---NTTISPATYKPNGEWELL	204	
Qy	200	GMPGKNTITIYACCPPEYVDVTFIMLRRTRLTYFFENLIVPCVLISWALLGFTLPDPSG	259	
Db	205	GLTSQRSIFFEYCCPEPYVDVTFIVSRRTRLTYGFNLLFPCMLISSALLSFTLPADCG	264	
Qy	260	EKUTLGVITLLSLTVFNLVAETLPQVSDAIPILGTYPNCIMFMVASSVVLTVVVLNYHH	319	
Db	265	EKINLGVITIFMSLCVFMIMVAEAMPQSDALPLTIQYFSCIMFQVGASVATVIALNFHH	324	
Qy	320	RTADIHER-MPQWIKSVFLQMLPWILRMRP-----GKKITRKTIMMNTMR	364	
Db	325	RSPEQYAPMKNKFLKTLGLMPLTLLGMRDPDVLSEVGHAYASDNKKKQYLIIEVERH	384	
Qy	365	ELEKERSKSLLANVLID-----DDFRHGPPPPNSTAST--GNLPGCCSIFRDFR	415	
Db	385	ILTRPENGNGHSADVDAVHLDLSTGNPHSDAKKSSPSKRTSASIMGWG-----	433	
Qy	416	RSFVRPSTMEDVGGGLGS-----	433	
Db	434	-----LPTT--QMGALDSSINKYTCTKVTRPLENGSATINHKSSPQINPINNNNIYKCAN	487	
Qy	434	-----HHREHLILIRELOFITARMKKADBEAELISDWKFAAMVVDFCLFVFTLFIIA	487	
Db	488	NQKTQPEDRFRPHILNELRVLISARVKEEAHHALQADMPASRVVDVRVCFIAFSFLPMC	547	
Qy	488	TVAVLLSAPHIIV	500	
Db	548	TAIISYNAPHLFV	560	

RESULT 11

A53956  
nicotinic acetylcholine receptor alpha-3 chain precursor, neuronal - human  
C;Species: Homo sapiens (man)  
C;date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 20-Aug-1999  
C;Accession: A53956; S21338  
R;Mihovilovic, M.; Roses, A.D.  
Exp. Neurol. 111, 175-180, 1991  
A;Title: Expression of mRNAs in human thymus coding for the alpha3 subunit of a neuronal  
A;Reference number: A53956; MUID:51114756; PMID:1989896  
A;Accession: A53956  
A;Status: preliminary  
A;Molecule type: mRNA

A;Residues: 1-503 <MI>  
A;Cross-references: GB:M37981; NID:g189252; PIDN:AAA59942.1; PID:g189253  
R;Anand, R.; Lindstrom, J.  
A;Submitted to the EMBL Data Library, June 1990  
A;Description: Nucleotide sequence of the mature human nicotinic acetylcholine receptor  
A;Reference number: S21338  
A;Accession: S21338  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 30-503 <ANA>  
A;Cross-references: EMBL:X53559; NID:g34985; PIDN:CAA37625.1; PID:g34986  
C;Genetics:  
A;Gene: GDB:CHRNA3  
A;Cross-references: GDB:125219; OMIM:118503  
A;Map position: 15q24-15q24  
C;Superfamily: acetylcholine receptor  
C;Keywords: neurotransmitter receptor

[illegible]

```

121  GTIQINVRSGSCLYVPGIIRKSTCKMDIANFPFDDQDCMKFGSWTYDGNQDLVLK 180
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
131  VDDKTRALLKYTGEVTWIPPAIFKSSCKIDVTFPPDYQNCMTKFGSWSYDKAKIDLVI 190
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181  DEAGGDSDFITNGEWWLIGMPGKKNTIITYACCEPYVDVTFIMIRRLTYFFNLIVP 240
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
191  G-SMMLKDWSEGEWAIIRKAPYKHDIKYNCCEIYDPDITYSLVIRRLPLFYINLIIP 249
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241  CVLIISSMALIGFTLPDPSGEKLTGLVTILASLVFLNLVAETLPQVSDAIPLLGYFNCI 300
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
250  CLLISFLTVLVPDSCGKVTLCISVLSLTVFLVAITETIPSTSLVPLIGEYLLFT 309
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301  MFWVASSVLTVVVLNVHHRBTADTHEMPOIKSVFLOLWFLMRSPGKKITRKTIMNN 360
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
310  MIFVTLISIVTVFVLNVHYATPTHTHMPWSWKVFLMLLPRVFMFTRP----- 357
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
361  TRMRELEKERSKSLIANVLDIDDDFRHGPPPNSTASTGNL-----GPGCS----- 408
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
358  -----TSNEGNAQKPRLYGAELSNLNCFSBAESKGCKEGYPC 395
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
409  -----IPTDRFRSFVRPSTMEDVG-GLGSHHRELHLILRELQFTARMKK 454
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
396  QDQMGCGYCHHRRIKISNFSANLTRSSSESVDVLSLSALSPEIKEAIOVKYIAENMKA 455
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
455  ADREAEALISDWKFAAMVVDRECLFVFTLFTIATVAVLL 493
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
456  QNEAKIQDDWKYVAMVIDRIFTLVVFTVLGILGTAGLFL 494
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12

A37040  
nicotinic acetylcholine receptor alpha-3 chain precursor, neuronal (version 2) - human  
C;Species: Homo sapiens (man)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 21-Jul-2000  
C;Accession: A37040; S24595  
R;Fornasari, D.; Chini, B.; Tarroni, P.; Clementi, F.  
Neurosci. Lett. 111, 351-356, 1990  
A;Title: Molecular cloning of human neuronal nicotinic receptor alpha-3-subunit.  
A;Reference number: A37040; MUID:90245296; PMID:2336208  
A;Accession: A37040  
A;Molecule type: mRNA  
A;Residues: 1-502 <FOR>  
A;Cross-references: EMBL:X52239; NID:q177897; PIDN:AAC84176.1; PID:q177898







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:18:28 ; Search time 9.35654 Seconds  
(without alignments)  
2788.120 Million cell updates/sec

Title: US-09-303-232-6

Perfect score: 2640

Sequence: 1 MAPMLAALALLPVSEQ.....LFTIATVALLSAPHIIQ 501

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1258.5	47.7	502	1 ACH7 HUMAN	P36544 homo sapien
2	1255.5	47.6	502	1 ACH7 MOUSE	P49582 mus musculus
3	1253	47.5	502	1 ACH7 CHICK	P22770 gallus gall
4	1246.5	47.2	502	1 ACH7 RAT	Q05941 rattus norv
5	1240.5	47.0	499	1 ACH7 BOVIN	P54131 bos taurus
6	1237.5	46.9	498	1 ACH1 CAEEL	P48180 caenorhabdi
7	989.5	37.5	503	1 ACH3 HUMAN	P32297 homo sapien
8	977.5	37.0	499	1 ACH3 RAT	P04757 rattus norv
9	970.5	36.8	495	1 ACH3 BOVIN	Q07263 bos taurus
10	964.5	36.5	567	1 ACH1 DROME	P09478 drosophila
11	948.5	35.9	496	1 ACH3 CHICK	P09481 gallus gall
12	946.5	35.9	496	1 ACH1 MANSE	P91766 manduca sex
13	944	35.8	557	1 ACH1 SCHGR	P23414 schistocerc
14	943	35.7	576	1 ACH2 DROME	P17644 drosophila
15	933	35.3	494	1 ACH6 HUMAN	Q15825 homo sapien
16	924.5	35.0	528	1 ACH2 CHICK	P09480 gallus gall
17	922.5	34.9	529	1 ACH2 HUMAN	Q15822 homo sapien
18	919.5	34.8	512	1 ACH3 CARAU	P18845 carassius a
19	917	34.7	511	1 ACH2 RAT	P12389 rattus norv
20	915.5	34.7	494	1 ACH6 CHICK	P49581 gallus gall
21	914.5	34.6	493	1 ACH6 RAT	P43143 rattus norv
22	910.5	34.5	495	1 ACHP RAT	P12392 rattus norv
23	908.5	34.4	519	1 ACHA DROME	P25162 drosophila
24	906	34.3	457	1 ACHA BOVIN	P02709 bos taurus
25	900	34.1	521	1 ACH3 DROME	P04755 drosophila
26	899.5	34.1	500	1 ACHN RAT	P12390 rattus norv
27	897.5	34.0	627	1 ACH4 HUMAN	P43681 homo sapien
28	896	33.9	456	1 ACHA CHICK	P09479 gallus gall
29	895	33.9	502	1 ACHN HUMAN	P17787 homo sapien
30	891	33.8	622	1 ACH4 CHICK	P09482 gallus gall
31	890.5	33.7	629	1 ACH4 MOUSE	O70174 mus musculus
32	890	33.7	457	1 ACHA RAT	P25108 rattus norv
33	890	33.7	470	1 ACHP CHICK	P26153 gallus gall

## RESULT 1

ID	ACH7_HUMAN	STANDARD;	PRT;	502 AA.
AC	P36544; Q15826; Q96RH2; Q99555; Q9BXH0;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Neuronal acetylcholine receptor protein, alpha-7 chain precursor.			
GN	CHRNA7 OR NACHRA7.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=94195283; PubMed=8145738;			
RA	Peng X., Katz M., Gerzanich V., Anand R., Lindstrom J.;			
RT	"Human alpha 7 acetylcholine receptor: cloning of the alpha 7 subunit from the SH-SY5Y cell line and determination of pharmacological properties of native receptors and functional alpha 7 homomers expressed in Xenopus oocytes.";			
RT	Mol. Pharmacol. 45:546-554(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Hippocampus;			
RA	Logel J., Drebing C., Barnhart M., Antle C., Leonard S.;			
RL	Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97062879; PubMed=8906617;			
RA	Elliot K.J., Ellis S.B., Berckhan K.J., Urrutia A.,			
RA	Chavez-Noriega L.E., Johnson E.C., Velicelebi G., Harpold M.M.;			
RT	"Comparative structure of human neuronal alpha 2-alpha 7 and beta 2-beta 4 nicotinic acetylcholine receptor subunits and functional expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and beta 4 subunits.";			
RL	J. Mol. Neurosci. 7:217-228(1996).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97162233; PubMed=9009220;			
RA	Groot Kormelink P.J., Luyten W.H.M.I.;			
RT	"Cloning and sequence of full-length cDNAs encoding the human neuronal nicotinic acetylcholine receptor (nAChR) subunits beta3 and beta4 and expression of seven nAChR subunits in the human neuroblastoma cell line SH-SY5Y and/or IMR-32.";			
RL	FEBS Lett. 400:309-314(1997).			
RN	[5]			
RP	REVISIONS.			
RA	Groot Kormelink P.J., Luyten W.H.M.I.;			
RL	Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Epidermal keratinocytes;			
RA	Arredondo J., Grando S.A.;			
RT	"Cloning cholinergic receptors in human keratinocytes.";			

34	887	33.6	498	1	ACHP_HUMAN	P30926 homo sapien
35	886	33.6	457	1	ACHA_MOUSE	P04756 mus musculus
36	885.5	33.5	459	1	ACHN_CARAU	P19370 carassius a
37	882.5	33.4	482	1	ACHA_HUMAN	P02708 homo sapien
38	881.5	33.4	461	1	ACHA_TORMA	P02711 torpado mar
39	881.5	33.4	630	1	ACH4_RAT	P09483 rattus norv
40	881	33.4	491	1	ACHN_CHICK	P09484 gallus gall
41	876.5	33.2	461	1	ACHA_TORCA	P02710 torpado cal
42	875.5	33.2	457	1	ACH2_XENLA	P05377 xenopus lae
43	869.5	32.9	457	1	ACH1_XENLA	P22456 xenopus lae
44	869	32.9	538	1	ACH8_CAEEL	P45963 caenorhabdi
45	863.5	32.7	456	1	ACHA_BRARE	Q98880 brachydanio

## ALIGNMENTS

Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
[7]  
SEQUENCE OF 17-502 FROM N.A.  
TISSUE=Brain;  
RA Doucette-Stamm L., Monteggia L.M., Donnelly-Roberts D., Wang M.T.,  
Lee J., Tian J., Giordano T.;  
RT "Cloning and sequence of the human  $\alpha 7$  nicotinic acetylcholine  
receptor.";  
EL Drug Dev. Res. 30:252-256(1993).  
RN [8]  
RP SEQUENCE OF 24-502 FROM N.A.  
TISSUE=Retina;  
RC MEDLINE=94245214; PubMed=8188270;  
RA Chini B., Raimondi E., Elgoyhen A.B., Moralli D., Balzaretto M.,  
Heinenann S.F.;  
RT "Molecular cloning and chromosomal localization of the human  $\alpha 7$ -  
nicotinic receptor subunit gene (CHRNA7).";  
RL Genomics 19:379-381(1994).  
RN [9]  
RP SEQUENCE OF 118-129 FROM N.A.  
MEDLINE=21818878; PubMed=11829490;  
RA Riley B., Williamson M., Collier D., Wilkie H., Makoff A.;  
RT "A 3-Mb map of a large segmental duplication overlapping the  $\alpha 7$ -  
nicotinic acetylcholine receptor gene (CHRNA7) at human 15q13-q14.";  
RL Genomics 79:197-209(2002).  
RN [10]  
RP MASS SPECTROMETRY.  
RC TISSUE=Breast cancer;  
RX MEDLINE=21829512; PubMed=11840567;  
RA Harris R.A., Yang A., Stein R.C., Lucy K., Brusten L., Herath A.,  
Parekh R., Waterfield M.D., O'Hare M.J., Neville M.A., Page M.J.,  
Zvelebil M.J.;  
RT "Cluster analysis of an extensive human breast cancer cell line  
protein expression map database.";  
RL Proteomics 2:212-223(2002).  
CC -!- FUNCTION: After binding acetylcholine, the ACHR responds by an  
extensive change in conformation that affects all subunits and  
leads to opening of an ion-conducting channel across the plasma  
membrane.  
CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-  
bungarotoxin. The structure is probably pentameric (By  
similarity).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- MASS SPECTROMETRY: MW=54157.68; METHOD=MALDI.  
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X70297; CAA49778.1; -;  
CC EMBL; U40583; AAA83561.1; -;  
CC EMBL; U62436; AAB40114.1; -;  
CC EMBL; Y08420; CAA69697.1; -;  
CC EMBL; AF385585; AAK68111.1; -;  
CC EMBL; L25827; -; NOT ANNOTATED CDS.  
CC EMBL; Z23141; CAA80672.1; -;  
CC EMBL; AF332758; AAK19515.1; -;  
CC PIR; G02259; G02259.  
CC PIR; I37185; ACHUA7.  
CC Genew; HGNC:1960; CHRNA7.  
CC MIM; 118511; -;  
CC GO; GO:0005892; C:nicotinic acetylcholine-gated receptor-chan. .; TAS.  
CC GO; GO:0015464; F:acetylcholine receptor activity; TAS.  
CC GO; GO:0004889; F:nicotinic acetylcholine-activated cation-se. .; TAS.  
CC GO; GO:000187; P:activation of MAPK; TAS.  
CC GO; GO:0006810; P:transport; TAS.  
CC InterPro; IPR006029; Neur\_chan\_memb.  
CC InterPro; IPR006202; Neur\_chan\_IIBD.

DR InterPro; IPR006201; Neur\_channel.  
DR Pfam; PF02931; Neur\_chan\_IIBD; 1.  
DR Pfam; PF02932; Neur\_chan\_memb; 1.  
DR PRINTS; PR00252; NRIONCHANNEL.  
DR TIGRFAMS; TIGR00860; LIC; 1.  
DR PROSITE; PS00236; NEUOTR\_ION\_CHANNEL; 1.  
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
KW Transmembrane; Multigene family.  
FT SIGNAL 1 22 BY SIMILARITY.  
FT CHAIN 23 502 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,  
FT ALPHA-7 CHAIN.  
FT DOMAIN 23 230 POTENTIAL.  
FT TRANSMEM 231 255 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 262 280 POTENTIAL.  
FT TRANSMEM 296 317 POTENTIAL.  
FT DOMAIN 318 469 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 470 490 POTENTIAL.  
FT DISULFID 150 164 BY SIMILARITY.  
FT DISULFID 212 213 ASSOCIATED WITH RECEPTOR ACTIVATION  
(BY SIMILARITY).  
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 90 90 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT TRANSMEM 133 133 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 11 11 A -> G (IN REF. 1 AND 7).  
FT CONFLICT 58 58 S -> N (IN REF. 2 AND 6).  
FT CONFLICT 134 134 S -> P (IN REF. 2 AND 6).  
FT CONFLICT 364 364 C -> S (IN REF. 8).  
FT CONFLICT 375 375 A -> G (IN REF. 1).  
FT CONFLICT 409 413 RMACS -> AMPAP (IN REF. 8).  
SQ SEQUENCE 502 AA; 56449 MW; D94B3A482EAA0E42 CRC64;  
Query Match 47.7%; Score 1258.5; DB 1; Length 502;  
Best Local Similarity 48.6%; Pred. No. 6.3e-102; Indels 51; Gaps 7;  
Matches 250; Conservative 75; Mismatches 138;  
Qy 8 LALLA-LLPVSEQGHEKRLNALLANYTLERFVANESEPLEYRFGTLTQOIIDVDKDN 66  
Db 10 LALAALLHVSLSQGEFQKLYELVKNYNPLERFVANDSQPLTVVFSLSLIQIMVDKDN 69  
Qy 67 QLLTNTIWSLEWYNDYLNWDSYGGVKDLRITPNKWKPDVLMYNSADSGFGTYGTN 126  
Db 70 QVLTNTIWLQMSWTDHYLQWNVSEYGVKTVFRFPDGGIILLYNSADSRFDFATFTN 129  
Qy 127 VVVRSGGCLVPPGIEFKSTCKMDIAWFPDQDCHDKMFGSWTYDGNQLDLVLKDEAGD 186  
Db 130 VLVNSSHCQVLPGLPKSSCYIDVRWFPDQVQCKLKFGSWSGWSLDLQMBE--AD 186  
Qy 187 LSDFITNGEWYLGMPGKKNITVACPEPYVDVFTTITMIRRTLYYFFNLIVPCVLIS 246  
Db 187 ISGYIPGEWDLVGIPIGRSERFYECKPEYDPDVTFTVWRRRTLYYGLNLIIPCVLISA 246  
Qy 247 MALLGFTLPDPSGKLTGLVITLLSLTVFLNVAETLPOVSDATPLLTGYENCIMFVVAS 306  
Db 247 LALLVFLPADSGEKISLIGITVLSLTVFLMLVAEIMPATSDSVFLTAQYFASITMIIVGL 306  
Qy 307 SVLTVVVLNVYHRTADITHMPOMIKSVFLQWLPIWLRMSRPGKKITRKTIMMTRREL 366  
Db 307 SVVTVIVLQYHHHDPGGKMPKWTRVILLNWCWAFLEMKRPGEDKVRPACQHKQRCSL 366  
Qy 367 ELKERSKSLIANVLDIDDDFRHGPPPNPNSTASTGNL-----GGPC 407  
Db 367 ASVEMSAVA-----PPP---ASGNLLYIGRGLDGVCHVCPTPDGSGVVC 407  
Qy 408 SIFTFDFRSFVRPSTMEDV--GGGLSGHREHLILRELQFTIARMKKADEAEELISDW 465  
Db 408 G-----RMACSTHDEHLHGQPEGDDLAKILEEVRYTIANRFRQCDESEAVCSEW 460  
Qy 466 KFAAMVDRFCLFVFTLTFTIATVAVLLSAPHII 499  
Db 461 KFAACVDRCLMAFSVFTIICTIGILMSAFNFV 494

RESULT 2

ACH7\_MOUSE  
ID ACH7\_MOUSE STANDARD; PRT; 502 AA.  
AC P49582;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.  
GN CHRNA7 OR ACR47.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c; TISSUE=Brain;  
RX MEDLINE=95324936; PubMed=7601470;  
RA Urtreger A., Seldin M.F., Baldini A., Beaudet A.L.;  
RT "Cloning and mapping of the mouse alpha 7-neuronal nicotinic  
acetylcholine receptor.";  
RL Genomics 26:399-402(1995).  
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an  
extensive change in conformation that affects all subunits and  
leads to opening of an ion-conducting channel across the plasma  
membrane.  
CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-  
bungarotoxin. The structure is probably pentameric (By  
similarity).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; L37663; AAC2053.1; -.  
DR PIR; A57175; A57175.  
DR MGD; MG1:99779; Chnra7.  
DR InterPro; IPR006029; Neur channel memb.  
DR InterPro; IPR006202; Neur\_chan\_LBD.  
DR InterPro; IPR006201; Neur\_chan.  
DR Pfam; PF02931; Neur\_chan\_LBD; 1.  
DR Pfam; PF02932; Neur\_chan\_memb; 1.  
DR PRINTS; PR00252; NRIONCHANNEL.  
DR TIGRfams; TIGR00860; LIC; 1.  
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.  
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
KW Transmembrane; Multigene family.  
FT SIGNAL 1 22 BY SIMILARITY.  
FT CHAIN 23 502 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,  
ALPHA-7 CHAIN.  
FT FT 23 502  
FT DOMAIN 23 230 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 231 255 POTENTIAL.  
FT TRANSMEM 262 280 POTENTIAL.  
FT TRANSMEM 296 317 POTENTIAL.  
FT DOMAIN 318 469 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 470 490 POTENTIAL.  
FT DISULFID 150 164 BY SIMILARITY.  
FT DISULFID 212 213 ASSOCIATED WITH RECEPTOR ACTIVATION  
(BY SIMILARITY).  
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 502 AA; 56631 MW; C9312E5226D120E3 CRC64;  
  
Query Match 47.6%; Score 1255.5; DB 1; Length 502;  
Best Local Similarity 49.6%; Pred. No. 1.2e-101;  
Matches 249; Conservative 78; Mismatches 148; Indels 27; Gaps 7;  
  
QY 8 LAL-LALLPVSQGPQHEKRLNALLANYTLERPVANSEPLEVRFLTLTQQIIDVDEN 66

Db 10 LALAAALHVSQGEFORLYKELVKNYPLERPVANDSOPLTVYFSLSLQIMDVDEKN 69  
QY 67 QLLITNIWLSLENDYNLRWDSSEYGGVKDLRTTPNKLMPDVMYNSADEGDTGYQTN 126  
Db 70 QVLTINILQMSWTDLHQLQNMSEYPGVKNVRPPDQGIWKPKDILLYNSADERDAFHTN 129  
QY 127 VVVRSGSGLYPPGIFKSTCKMDIAWPFDDHCHDKMGKSGSWYDGNQLDLVKDAGGD 186  
Db 130 VLVNASGHQYLPPIGFKSCYIDVRWFFDVQOQCKLFGSWSYGGMSLDLQMQE---AD 186  
QY 187 LSDFTNGSWYLLGMPKKNITTYACCPREYVDVTFIMIRRTLYFFENLIYPCVLISS 246  
Db 187 ISSYIENGWDLMGIFPKRNEKFECKEPPDVTIVIMRRRLTYGLNLLPCVLISA 246  
QY 247 MALLGFTLPDSCGKLTGLVTILLSTLVAETLPOVSDAIPLLGYFNCIMFMVAS 306  
Db 247 LALLVLLPADSGEKISLGTIVLLSTVFWMLLVAEIMPATSDSVPLIAQVFASTMIIVGL 306  
QY 307 SVVLTIVVLYNHRTADIHEMPQWIKSVFLOMPLWILRMRSPGKTKTRTIMMNTMREL 366  
Db 307 SVVTVIVLYRHHDDPGGKMPKWTIRILLNWCAMFLRMKRPGEKVRPACQHKPRCSL 366  
QY 367 ELKERSS---KSLLANVLIDDDRR-----HGPPPNSTASTGNLGPCCSIFRTDFRS 417  
Db 367 ASVELSAGAGPPTSNGNLLYI--GFRGLEGMHCAPTPDSGVVCGRL--ACSPHDEHLMH 422  
QY 418 FVRPSTMEDVGGGLGSHRHLHLRELQFITARMKKADDEAEELISDWKFAAMVVDRCFL 477  
Db 423 GTHPSDGP-----DLAKILEEVYIANRRCQDESEVICSEWKEFAACVVDRLCL 472  
QY 478 FVPTLTITATVAVLLSAPHII 499  
Db 473 MAPSVETIITIGILMSAPNFV 494  
  
RESULT 3  
ACH7\_CHICK STANDARD; PRT; 502 AA.  
AC P22770;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=91097796; PubMed=1702646;  
RA Couturier S., Bertrand D., Matter J.M., Hernandez M.C., Bertrand S.,  
Millar N., Valera S., Barkas T., Ballivet M.;  
RT "A neuronal nicotinic acetylcholine receptor subunit (alpha 7) is  
developmentally regulated and forms a homo-oligomeric channel blocked  
by alpha-BTX.";  
RL Neuron 5:847-856(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=90315158; PubMed=2369519;  
RA Schoeffer R., Conroy W.G., Whiting P., Gore M., Lindstroem J.;  
RT "Brain alpha-bungarotoxin binding protein cDNAs and MAbS reveal  
subtypes of this branch of the ligand-gated ion channel gene  
superfamily.";  
RL Neuron 5:35-48(1990).  
RN [3]  
RP SEQUENCE OF 1-18 FROM N.A.  
RC STRAIN=White leghorn; TISSUE=Erythrocyte;  
RX MEDLINE=93049204; PubMed=1425587;  
RA Matter-Sadzinski L., Hernandez M.C., Roztocil T., Ballivet M.,

RA Matter J.M.;  
 RT "Neuronal specificity of the alpha 7 nicotinic acetylcholine receptor  
 RT promoter develops during morphogenesis of the central nervous  
 RT system.";  
 RL EMBO J. 11:4529-4538(1992).  
 RN [4]  
 RN SEQUENCE OF 24-47.  
 RP TISSUE=Brain;  
 RX MEDLINE=85270494; PubMed=3860855;  
 RA Conti-Tronconi B.M., Dunn S.M.J., Barnard E.A., Dolly J.O., Lai F.A.,  
 RA Ray N., Raftery M.A.;  
 RT "Brain and muscle nicotinic acetylcholine receptors are different but  
 RT homologous proteins.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:5208-5212(1985).  
 RN [5]  
 RN MUTAGENESIS OF LEU-270.  
 RX MEDLINE=92049732; PubMed=1719423;  
 RA Recah F., Bertrand D., Galzi J.-L., Devillers-Thiery A., Malle C.,  
 RA Hussy N., Bertrand S., Ballivet M., Changeux J.-P.;  
 RT "Mutations in the channel domain alter desensitization of a neuronal  
 RT nicotinic receptor";  
 RL Nature 353:846-849(1991).  
 RN [6]  
 RN MUTAGENESIS TO CONVERT ION SELECTIVITY FROM CATIONIC TO ANIONIC.  
 RX MEDLINE=93024917; PubMed=1383829;  
 RA Galzi J.-L., Devillers-Thiery A., Hussy N., Bertrand S.,  
 RA Changeux J.-P., Bertrand D.;  
 RT "Mutations in the channel domain of a neuronal nicotinic receptor  
 RT convert ion selectivity from cationic to anionic.";  
 RL Nature 359:500-505(1992).  
 CC -!- FUNCTION: After binding acetylcholine, the ACHR responds by an  
 CC extensive change in conformation that affects all subunits and  
 CC leads to opening of an ion-conducting channel across the plasma  
 CC membrane.  
 CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-  
 CC bungarotoxin. The structure is probably pentameric (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- DEVELOPMENTAL STAGE: Alpha-7 transcripts transiently accumulate  
 CC in the developing optic tectum between E5 and E16.  
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X52295; CAA36543.1; -;  
 CC EMBL; X68246; CAA48317.1; -;  
 CC EMBL; X68586; CAA48576.1; -;  
 CC PIR; JN0113; JN0113.  
 CC PDB; 1KC4; 17-APR-02.  
 CC PDB; 1KL8; 17-APR-02.  
 CC InterPro; IPR006029; Neu channel memb.  
 CC InterPro; IPR006202; Neu channel LED.  
 CC InterPro; IPR006201; Neu channel.  
 CC Pfam; PF02931; Neu chan LBD; 1.  
 CC Pfam; PF02932; Neu chan memb; 1.  
 CC PRINTS; PR00252; NRIONCHANNEL.  
 CC TIGR; TIGR00860; LIC; 1.  
 CC PROSITE; PS00236; NEUROTR ION CHANNEL; 1.  
 KW Postsynaptic membrane, Ionic channel, Glycoprotein; Signal;  
 KW Transmembrane, Multigene family; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 502 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,  
 FT ALPHA-7 CHAIN.  
 FT DOMAIN 24 230 EXTRACELLULAR.  
 FT TRANSMEM 231 255  
 FT TRANSMEM 262 280  
 FT TRANSMEM 296 317

FT DOMAIN 318 469 CYTOPLASMIC.  
 FT TRANSMEM 470 490  
 FT DISULFID 150 164 BY SIMILARITY.  
 FT DISULFID 212 213 ASSOCIATED WITH RECEPTOR ACTIVATION  
 FT (BY SIMILARITY).  
 FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT MUTAGEN 270 270 L-S,S,T: SUPPRESSES INHIBITION BY THE  
 FT OPEN-CHANNEL BLOCKER QX-222.  
 FT CONFLICT 26 27 QR -> ET (IN REF. 3).  
 SQ SEQUENCE 502 AA; 56946 MW; 572325D4309AD2FD CRC64;  
 Query Match 47.5%; Score 1253; DB 1; Length 502;  
 Best Local Similarity 48.8%; Pred. No. 1.9e-101;  
 Matches 250; Conservative 81; Mismatches 145; Indels 36; Gaps 8;  
 Qy 2 APMALALLALLLPVSEQGEHEKRLNALLNANTYLERPVANSEPLEVRFGLTQQIID 61  
 Db 5 ALMLWLLAAAGLVRESLQGEFQKLYKELLKNYPLERPVANDSQPLTVVFTLSLMQIMD 64  
 Qy 62 VDEKNQLLTINILSLEWNYLNWDSYGGVUKDLRITPNKLPKDPVLVMSADEGFDG 121  
 Db 65 VDEKNQVLTINILQWYTDHYLQWNVSEYPGVKNVFPDGLIWKPDILLYNSADERFDA 124  
 Qy 122 TYQTVVVRSGSLVYVPGIEFKSTCKMDIAWPFDDQHKDMKFGSWTYDGNQLDLVKD 181  
 Db 125 TETNVLVNSGHQCQLPPGLFKSSCYDVRWPFDDVQKCNKFGSWTYGWSLDLQMQE 184  
 Qy 182 EAGGDLSDFTINGEWYILGMPGKNTITYACCPPEYVDVFTIMIRRTLYFFNLIVPC 241  
 Db 185 --ADISSYISNGEWDLVGIPGRKTESFYECCKEYPDITFTVTMRRTLYVGLNLLIPC 241  
 Qy 242 VLTISSMALLGFTLPDSDGKLTIGVTILLSITVFLNLVAETLPQVSDAIPLLGYFNCIM 301  
 Db 242 VLISALALLVFLLPADSGEKISLITVLLSLTVFMLLVAEIMPATSDSVPLIAQYFASMT 301  
 Qy 302 FMVASSVVLTVVLYNHHRTADTHEMPQWIKSVFLQWLPILTRMSRPGKTKTRTIMNT 361  
 Db 302 IIVGLSVVTVVLYQYHHHPDGGKMPKTRVILLNWCAMFLMKRPGEDKVRPACQHKQ 361  
 Qy 362 RMRELEKERSKSKS-----LLANVLIDIDDDFR-----HGPPPNSTASTGNLGPCCSIFRT 412  
 Db 362 RRCSSLSMWNVTSGQCCSNGNMLYI--GFRGLDGVHCTPTTDSGVICGRM--TCS----- 413  
 Qy 413 DFRSRVFRPTMEDVGGGLSHH-----RELHLILRELOPITARMKKADEAEILSDWKF 467  
 Db 414 -----PTEENL--LHSGHPSEGDPLAKILEEVRYIANFRDQDEEAICNEMKF 462  
 Qy 468 AAMVWDRFCLFVFTLTFTIATVAVLLSAPHII 499  
 Db 463 AASVVDRLCLMAFSVFIICTIGILMSAPNFV 494  
 RESULT 4  
 ACH7 RAT  
 ID ACH7 RAT STANDARD; PRT; 502 AA.  
 AC Q05941;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 GN Neuronal acetylcholine receptor protein, alpha-7 chain precursor.  
 DE CHRNA7 OR ACRA7.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=93147931; PubMed=7678857;  
 RA Seguela P., Wadiche J., Dineley-Willer K., Dani J.A., Patrick J.W.;  
 RT "Molecular cloning, functional properties, and distribution of rat

brain alpha 7: a nicotinic cation channel highly permeable to calcium.";

J. Neurosci. 13:596-604 (1993).

[2]

SEQUENCE FROM N.A.

STRAIN=Sprague-Dawley; TISSUE=Brain;

RA Boulter J.;

Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.

[3]

REVISION TO 363..

RC STRAIN=Sprague-Dawley; TISSUE=Brain;

RA Hartley M.;

Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an

extensive change in conformation that affects all subunits and

leads to opening of an ion-conducting channel across the plasma

CC membrane.

CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-

bungarotoxin. The structure is probably pentameric (By

CC similarity).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

at the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; S53987; AAB25224.2; -;

DR EMBL; L31619; AAC33136.1; -;

DR PIR; T01378; T01378.

DR InterPro; IPR006029; Neu channel memb.

DR InterPro; IPR006202; Neur chan IBD.

DR InterPro; IPR006201; Neur channel.

DR Pfam; PF02931; Neur\_chan\_IBD; 1.

DR Pfam; PF02932; Neur\_chan\_memb; 1.

DR PRINTS; PR00252; NRIONCHANNEL.

DR TIGRFAMs; TIGR00860; LIC; 1.

DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.

KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;

Transmembrane; Multigene family.

FT SIGNAL 1 22 BY SIMILARITY.

FT CHAIN 23 502 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,

ALPHA-7 CHAIN.

FT DOMAIN 23 230 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 231 255 POTENTIAL.

FT TRANSMEM 262 280 POTENTIAL.

FT TRANSMEM 296 317 POTENTIAL.

FT DOMAIN 318 469 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 470 490 POTENTIAL.

FT DISULFID 150 164 BY SIMILARITY.

FT DISULFID 212 213 ASSOCIATED WITH RECEPTOR ACTIVATION

(BY SIMILARITY).

FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 447 447 N -> F (IN REF. 2).

FT CONFLICT 469 469 P -> R (IN REF. 2).

SQ SEQUENCE 502 AA; 56410 MW; 00996E74EC7B9A56 CRC64;

Query Match

Best Local Similarity 47.2%; Score 1246.5; DB 1; Length 502;

Matches 248; Conservative 78; Mismatches 149; Indels 27; Gaps 7;

QY 8 LAL-LALLPVSEGGPEKRLNALLANYTLERPVANSEPLEVRGLFIQIIVDEKN 66

Db LALAAALLHVSLQGEFQRRLYKELVKNYPLERPVANDSQPLTVYFSLQLQIMVDEKN 69

QY 67 QLLITNIWLSLENDYNLRWNSDSEYGGVKDLRTTPNKLKPDVLMVNSADEGFGTYQTN 126

Db 70 QVLTNIWLSWTDHYLQWNMSEYGVKNVRFPGQIWKPDILLNSADERDATEHTN 129

QY 127 VVVRSGSCLYVPPGIFKSTCKMDIAWFFPDQHQCKMFGSWTYDGNQLDLVLKAGGD 186

Db 130 VLVNASHGCHQYLPPIGIFKSSCYDVRWFFPDVQCKLKFGSWYGGWSLDLQWQE--AD 186

QY 187 LSDFTINGEWYLLGMPGKNTIYACCPYVDVVTFTIMRRRTLYFFNLIYPCVLISS 246

Db 187 ISSYIENGWDLMGIFGKNEKFECCBYPDVTYVTMMRRRTLYGLNLLPCVLISA 246

QY 247 MALLGFTLPDSGEKLTGVTLLSLTVFLNLVAETLPQVSDAIPLLTGVENCIMFMVAS 306

Db 247 LALLVLLPADSGEKISLGTVLLSLTVFLNLVAETLPQVSDAIPLLTGVENCIMFMVAS 306

QY 307 SVLTVVVLNYHRTADIHEMPOWIKSVFLQWLPWILRMSRPGKTKTRKTMNTRREL 366

Db 307 SVVTVIVLRYHHDPDGGKMPKWTRIILLNWCWFLRMKRPGEKVRFACQKPRCSL 366

QY 367 ELKERSS---KSLLANVLIDDDFR-----HGPPPPNSTAGNLPGGCSIFRTDFRS 417

Db 367 ASVELSAGAGPPTSGNMLIY--GFRLEGMHCAPTDPSGVVCGRL--ACSPTHDEHLMH 422

QY 418 FVRPSTMEDVGGGLGSHHRELHLILRELOFITARMKKADBEAELISDWKFAAMVDFRCL 477

Db 423 GHPSDGD-----DLAKILEEVYIANRRCQDESEVICSEWKFACVVDPLCL 472

QY 478 FVFTFTIATVAVLLSAPHII 499

Db 473 MAPSVFTIITIGILMSAPNFV 494

RESULT 5

ACH7\_BOVIN

ID ACH7\_BOVIN STANDARD; PRT; 499 AA.

AC P54131;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.

GN CHRNA7.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).

RC TISSUE=Adrenal medulla;

RX MEDLINE=95346009; PubMed=7620615;

RA Garcia-Guzman M., Sala F., Sala S., Campos-Caro A., Stuehmer W.,

RA Gutierrez L., Criado M.;

RT "Alpha-Bungarotoxin-sensitive nicotinic receptors on bovine

chromaffin cells: molecular cloning, functional expression and

RT alternative splicing of the alpha 7 subunit.";

RL Eur. J. Neurosci. 7:647-655 (1995).

CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an

extensive change in conformation that affects all subunits and

leads to opening of an ion-conducting channel across the plasma

CC membrane.

CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-

bungarotoxin. The structure is probably pentameric (By

similarity). Homo-oligomer of the short form gives rise to

CC unfunctional channels, as does coexpression of both long and short

CC forms of the receptor.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=Long;

CC IsoID=P54131-1; Sequence=Displayed;

CC Name=Short;

CC IsoID=P54131-2; Sequence=VSP 000075;

CC -!- TISSUE SPECIFICITY: At least in chromaffin cells.

CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.











DR TIGRFAMS; TIGR00860; LIC; 1.  
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.  
KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
KW Transmembrane; Multigene family.  
FT SIGNAL 1 21 POTENTIAL.  
FT CHAIN 22 495 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,  
FT DOMAIN 22 230 ALPHA-3 CHAIN.  
FT TRANSMEM 231 255 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 263 281 POTENTIAL.  
FT TRANSMEM 296 318 POTENTIAL.  
FT TRANSMEM 319 467 CYTOPLASMIC (BY SIMILARITY).  
FT TRANSMEM 468 487 POTENTIAL.  
FT DISULFID 149 163 BY SIMILARITY.  
FT DISULFID 213 214 ASSOCIATED WITH RECEPTOR ACTIVATION  
FT CARBOHYD 45 45 (BY SIMILARITY).  
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 495 AA; 56914 MW; 322825629821EA07 CRC64;  
Query Match 36.8%; Score 970.5; DB 1; Length 495;  
Best Local Similarity 39.8%; Pred. No. 8.3e-77;  
Matches 198; Conservative 93; Mismatches 181; Indels 25; Gaps 7;  
QY 5 LAALALLLPVSEQGPHEKRLNALLANVNTLPRVANESEPLEVRFGLTLOQIIDVDE 64  
DB 7 LRLRLRLRLPVLASTDAEHLRPERLPEDYNIIRPVANVSDPVIQFEVMSQLVKVDE 66  
QY 65 KNOLLITNLSLEWMDYNLRWDSYGVKDLRITNKLKPKDVLVWNSADEGFDGTQY 124  
DB 67 VQIMETNLWLKQWINDYKLNWPSDYDGAEFMVPKWKPDVLYNNVAGVDFQVDDK 126  
QY 125 TNVVRGSGSLYPPGIFKSTKMDIAWPFDDQHCMDKFGSWTYDGNOLDVLKDBAG 184  
DB 127 TKALLKTYGEVTPWIPALPKSSCIDYTPFDYQNTCKMFGSWYDKAKIDVLIG-SS 185  
QY 185 GGLSDFTINGEWLIGMPGKNTITYACCPPEYVDVFTTIRRRITLYFNFNLVPCVLI 244  
DB 186 MNLKDYWESGEWAIKAPGYKHDIKYCNCEIYDITVSLVIRLPLFYTNLIIPCLLI 245  
QY 245 SSMAILGFTLPPDSEKLTGLVTILLSLTVFLNVAETLQVSDAIBLGTIFYNCIMFW 304  
DB 246 SFUTVLVFLYLPSCGKVKTCISVLSLTVFLVITETIPSTSLVILIGEYLLFTMIFV 305  
QY 305 ASSVVLTVVNLVYHRTADITHMPQWIKSVFLQWLPILMRSPGK---KITRKTIMWNT 361  
DB 306 TLSIVITVFLVNLVHYRTPTHMPWVKTFILNLLPRVMTFRPASNEGTVQRPFPYSA 365  
QY 362 RMRELEKRSKSLLANVLIDDDFRHGGPPPNSTAGNLGPGCSIFR---TDPRRSF 418  
DB 366 ELSNLNCFSRIESKVC-----KEGYP-----CQDGLCG-YCHHRRAKISNFSANL 409  
QY 419 VRPSTMEDVGG--GLGSHRELHLILRELQFITARMKKADEAEELISDMKFAAMVDRFC 476  
DB 410 TRSSSESVDAVLSLSALSPEIKAEISQVKYIAENMKQAQNEAKIEIQDDWKYVAMVIDRIF 469  
QY 477 LFVFTLTIATVAVLL 493  
DB 470 LWVFILVILGTAGLFL 486  
RESULT 10  
ACHI DROME  
ID -ACHI DROME STANDARD; PRT; 567 AA.  
AC P09478; OSVC74;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Acetylcholine receptor protein, alpha-like chain 1 precursor.  
GN NACR-ALPHA-96AA OR ACRB OR ALS OR ACR96AA OR CG5610.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Oregon-R;  
RX MEDLINE=88283626; PubMed=2840281;  
RA Bossy B., Ballivet M., Spierer P.;  
RT "Conservation of neural nicotinic acetylcholine receptors from  
RL Drosophila to vertebrate central nervous systems.";  
EMBO J. 7:611-618(1988).  
[2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Berkley;  
RX MEDLINE=20196006; PubMed=107311132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Jastoli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclab J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
[3]  
RN REVISIONS.  
RP MEDLINE=22426069; PubMed=12537572;  
RX Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kinkner J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RL systematic review";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an  
CC extensive change in conformation that affects all subunits and  
CC leads to opening of an ion-conducting channel across the plasma  
CC membrane.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- TISSUE SPECIFICITY: CNS in embryos.  
CC -!- DEVELOPMENTAL STAGE: Late embryonic, late pupal and second instar  
CC larvae stages.



```
DR pfam; PF02931; Neur chan LBD; 1.
DR pfam; PF02932; Neur chan memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMS; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEURCTR_ION_CHANNEL; 1.
KW postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family.
FT SIGNAL 1 22
FT CHAIN 23 496
FT NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
FT ALPHA-3 CHAIN.
FT EXTRACELLULAR.
FT DOMAIN 23 231
FT TRANSMEM 232 256
FT TRANSMEM 264 282
FT TRANSMEM 298 319
FT DOMAIN 320 468
FT TRANSMEM 469 488
FT DISULFID 150 164
FT DISULFID 214 215
FT CARBOHYD 46 46
FT CARBOHYD 163 163
FT SEQUENCE 496 AA; 57027 MW; PD25BC02A9B601FF CRC64;
Query Match 35.9%; Score 948.5; DB 1; Length 496;
Best Local Similarity 37.7%; Pred. No. 6.9e-75;
Matches 195; Conservative 89; Mismatches 168; Indels 65; Gaps 6;
QY 7 ALALLALLPVSEOG-----PHEKRLNALLANNYTLERPVANESSEPLEVRFGLTQQIIDV 62
Db 6 ALLTTAAVCILFQCGGSEPHRIYAALFRNYOFVRPVKNASDEVITQFEVSMQVKV 65
QY 63 DEKNQLATITNLSLEWYNDYLRWNSYGVGKDLRITPNKLPKPDVLMYNSADEGFGDT 122
Db 66 DEVNQIMETNLWLXHWNDYKLRWNPVDYGGAEIRVPSGQIWKPDIVLYNNAVGDFQVD 125
QY 123 YQTNVVRSGGSCLYVPPGIPKSTCKMDIAWPPDDQCHDMKFGSWYDGNQLDLVLKDE 182
Db 126 DKTKALLKYTGDTWTPPAIFKSCKIDVTYFPEDYQNCITMKFGSWSDKAKIDLVLIGS 185
QY 183 AGGDLSPITNGEVLIGMCPGKNTITVACCPPEVPVDVTETIMIRRLTYFFENLIPCV 242
Db 186 T-MNLKDYWESGEWAILKAPGYKDIKYNCCSEIYTDITYSLYIRRLPLFTIINMIPCL 244
QY 243 LISSMALIGFTLPDSEKLTGLVTILSLVFLNVAETLPQVSDAIPLLGTGYFCIMF 302
Db 245 LISFLTVLVFLPSDQEKVTLCSVLSLVFLVITETIPSTSLVPLIGEVLLFTMI 304
QY 303 MVASSVLTVVVLNVHHTADIHEMPQWIKSVFLQWLPLWILRMRPGKKITRKTMMNTR 362
Db 305 FVTLISIVTVFLNVHVHTPTKTHMPVWVRTIFLNLPRIMFMTRP----- 350
QY 363 MRELELKERSKSLANVLDDDFRHPGPPNPSTASGNLG----- 404
Db 351 -----TSDENNQKPFYTFSEFNSLNCNFSSEIKCKDGFGVQCD 390
QY 405 PGCSIFR-----TDFRRSFVRPSTMEDVGG--GLGSHHRELHLILRBLQFITARMKAD 456
Db 391 MACSCCQYQRMKFSDFSGNLTRSSSESVDPLFSFVLSPEMRDAIESVKVIAENMQQN 450
QY 457 BEALISDWKFAAMVVRFCFLFVTLFTIATVAVLL 493
Db 451 EAKGIQDDWKYVAMVIDRIFLWVFLVLCILGTAGLFL 487
RESULT 12
ACHI MANSE
ID ACHI MANSE STANDARD; PRT; 516 AA.
AC P91766;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acetylcholine receptor protein, alpha-like chain precursor (MARAL).
GN ARAL.
```

```

Db 203 SVWDILVPAVNEKFEYCCDEPYLDITENITMRKILFYVNLIPCMGISFLVIVF 262
QY 253 TLPDSEGEKUTLGVITLLSTVFNLVAETLPQVSDAIPLLGYFNCIMFMVASSVLTV 312
Db 263 YLPDSGEKVSLSISLLSTVFLLAEIIPPTSIVPLLGKFLVFTMLDTRFSICVTV 322
QY 313 VVLNYHRTADIHEMPQWIKSVFLOLWPILWRSPGKKI---TRKTMNTRMRELEIX 369
Db 323 VVLNVHFRSQTTMSPWRRVFIHVLPRLVWRPHYRDLPHRSRPAGLVTGAGETTLW 382
QY 370 ERSKSLANVLIDDIFRGCPPPPNSTASTGNLGP--GCSIFRTDPRRSFVRPSTMEDV 427
Db 383 DEGSPGV-----PAPRPPCAPPLAPCAACA-----PAEPAL 416
QY 428 GGGLGSHR--ELHLILRELQFTARMKKADEABELISDWKFAAMVVDRCLVFVLTFTI 485
Db 417 CDALRRWHRCPELHKAIDGINIYADQTRKEESTRVKEDWKYVAMVLDRLPFLIFTAVV 476
QY 486 IATVAVLSAP 496
Db 477 VGSAGIILQAP 487

RESULT 13
ACHL_SCHGR
ID ACHL_SCHGR STANDARD; PRT; 557 AA.
AC P23414;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acetylcholine receptor protein, alpha-l1 chain precursor.
OS Schistocerca gregaria (Desert locust).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthoptera; Acrididae; Orthoptera; Caelifera; Acridomorpha;
OC Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
OX NCBI_TaxID=7010;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9102263; PubMed=1702381;
RA Marshall J., Buckingham S.D., Singai R., Lunt G.G., Goosey M.W.,
RA Barlison M.G., Sattelle D.B., Barnard E.A.;
RT "Sequence and functional expression of a single alpha subunit of an
RT insect nicotinic acetylcholine receptor.";
RL EMBO J. 9:4391-4398(1990).
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X55439; CAA39081.1; -.
CC PIR; S12359; S12359.
CC InterPro; IPR006023; Neu channel memb.
CC InterPro; IPR006202; Neu_chan_LBD.
CC InterPro; IPR006201; Neu channel.
CC Pfam; PF02931; Neu chan IBD; 1.
CC Pfam; PF02932; Neu chan memb; 1.
CC PRINTS; PR00252; NEIONCHANNEL.
CC TIGRfam; TIGR00860; LIC; 1.
CC PROSITE; PS00236; NEUROTR ION_CHANNEL; 1.
CC Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
CC Transmembrane; Multigene family.
CC SIGNAL 1 23 POTENTIAL.

```

```

FT CHAIN 24 557 ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA-L1
FT CHAIN.
FT EXTRACELLULAR.
FT DOMAIN 24 244
FT TRANSMEM 245 266
FT TRANSMEM 274 294
FT TRANSMEM 308 329
FT DOMAIN 330 500
FT TRANSMEM 501 523
FT DISULFID 151 165
FT DISULFID 224 225
FT CARBOHYD 47 47
FT CARBOHYD 235 235
FT DOMAIN 382 400
FT DOMAIN 406 422
SQ SEQUENCE 557 AA; 63026 MW; 168389C87DFDF3E CRC64;

Query Match 35.8%; Score 944; DB 1; Length 557;
Best Local Similarity 36.9%; Pred. No. 2e-74;
Matches 201; Conservative 96; Mismatches 181; Indels 66; Gaps 7;

QY 1 MAPMLAALALLPVSEOGPEKRLNLNALLANYTLEREVANESPEFVRFGLTQQII 60
Db 5 LPPMLLLLLLLLHPAAANPDAKRLVDLLSNYNRLRPVSNNNTDVLVKGLRLSQLI 64
QY 61 DVDEKQQLITNIWLSLEWYNDYNLRWNSHYGGVKDLRITPNKLPDVLMYNSADGFD 120
Db 65 DLNLKQILITNVWLEHWDPAHYGGVTELYPSEHILWLPDILVYNNAGEYV 124
QY 121 GTYTNNVVRSGSCLYVPGIFKSTCKMDIAFPDQDQCHDKMFGSWTYDGNQDLIV-- 178
Db 125 VTTMTXAVLHHTKGVVTPPAIFKSCCEIDRVFPDQQTCTFMKFGSWTYDGGQIDLKHI 184
QY 179 -----LKDEAGGLSDFITNGEWYLGMPGKNTITYACCPYFYVDVTFIMIRRTLY 232
Db 185 NQKYVDNKVKVGIDIREYVPSVEWDILGVPAERHEKYVPCCAEPYDIFENILRRKTLF 244
QY 233 YFFNLIVPCVLISSMALLGFTLPPDSGEKUTLGVITLLSTVFNLVAETLPQVSDAIP 292
Db 245 YTVNLIVPCVCGISYLSVLVFPADSGEKIALCISILLSTQTMFLLISEIPSTSLAPL 304
QY 293 LGTYFCIMFMVASSVLTVVVLYNHYRTADIHEMPQWIKSVFLOLWPILWRSPGK-- 350
Db 305 LGXYLLFTVVLGSLVVTIMVNLNHYRKPSTHKMAPWVKVFIKRLPKLLMRVBEQLL 364
QY 351 -KITRKTIMMTRRELKERSKSLA--NVLDIDDDFRH----- 389
Db 365 ADLASKRLLRHAHNSKLSAAAAAASSSAASPSLSLHHHLHQHQLHLQLHLQ 424
QY 390 GPPPPNST-ASTGNLPGGCSIPTDPRFSVRPSTMEDVGGGLGS----- 433
Db 425 RPGGCNGLHSATNRFSGSAGAF-----GGLPVVGLDGSLDVATR 466
QY 434 -HREHLRLTLRELQFTARMKKADEABELISDWKFAAMVVDRCLVFVLTFTIATVAVL 492
Db 467 KYPFELEKAHNVLFIONHQRODEDAEDQDWGFVAMVLDRLFLWIFTIASIVGTFAIL 526
QY 493 LQAP 496
Db 527 CEAP 530

RESULT 14
ACH2_DROME
ID ACH2_DROME STANDARD; PRT; 576 AA.
AC P17644; Q9VC73;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Acetylcholine receptor protein, alpha-like chain 2 precursor.
GN NACR-ALPHA-96AB OR ACR96 OR SAD OR ACR96AB OR CG6844.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Head;  
RX MEDLINE=20301489; PubMed=2114015;  
RA Baumann A., Jonas P., Gundelfinger E.D.;  
RT "Sequence of D alpha 2, a novel alpha-like subunit of Drosophila  
RT nicotinic acetylcholine receptors.";  
RL Nucleic Acids Res. 18:3640-3640(1990).  
[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Head;  
RX MEDLINE=90353591; PubMed=2117557;  
RA Jonas P., Baumann A., Merz B., Gundelfinger E.D.;  
RT "Structure and developmental expression of the D alpha 2 gene  
RT encoding a novel nicotinic acetylcholine receptor protein of  
RT Drosophila melanogaster.";  
RL FEBS Lett. 269:264-268(1990).  
[3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Head;  
RX MEDLINE=90360975; PubMed=1697262;  
RA Sawruk E., Schloss P., Betz H., Schmitt B.;  
RT "Heterogeneity of Drosophila nicotinic acetylcholine receptors: SAD,  
RT a novel developmentally regulated alpha-subunit.";  
RL EMBO J. 9:2671-2677(1990).  
[4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RA MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang X., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,  
RA Swirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-X., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
[5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley; TISSUE=Head;  
RX MEDLINE=22426066; PubMed=12537569;

RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
RA George R.A., Guarin H., Krommiller B., Pacle J.M., Park S., Wan K.H.,  
RA Rubin G.M., Celniker S.E.;  
RT "A Drosophila full-length cDNA resource.";  
CC Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an  
CC extensive change in conformation that affects all subunits and  
CC leads to opening of an ion-conducting channel across the plasma  
CC membrane.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- TISSUE SPECIFICITY: CNS in embryos.  
CC -!- DEVELOPMENTAL STAGE: Late embryonic and late pupal stages.  
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; X52274; CAA36517.1; -;  
CC EMBL; X53583; CAA37652.1; -;  
CC EMBL; AE003748; AAF56303.1; -;  
CC EMBL; AY058446; AAL13675.1; -;  
CC F1R; S11679; ACEFA2.  
CC FlyBase; FBgn0000039; nAChR-alpha-96Ab.  
CC InterPro; IPR006029; Neu\_chan memb.  
CC InterPro; IPR006202; Neu\_chan LBD.  
CC InterPro; IPR006201; Neu\_chan.  
CC Pfam; PF02931; Neur\_chan\_LBD; 1.  
CC Pfam; PF02932; Neur\_chan\_memb; 1.  
CC PRINTS; PR00252; NRIONCHANNEL.  
CC TIGRFAMS; TIGR00860; LIC; 1.  
CC PROSITE; PS00236; NEUOTR ION CHANNEL; 1.  
CC Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
CC Transmembrane; Multigene family.  
CC SIGNAL 1 21 PROBABLE.  
CC CHAIN 22 576 ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA-  
CC LIKE CHAIN 2.  
CC DOMAIN 22 261 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 262 285 POTENTIAL.  
CC TRANSMEM 293 311 POTENTIAL.  
CC TRANSMEM 327 346 POTENTIAL.  
CC DOMAIN 347 526 CYTOPLASMIC (POTENTIAL).  
CC TRANSMEM 527 545 POTENTIAL.  
CC DISULFID 169 183 BY SIMILARITY.  
CC DISULFID 243 244 ASSOCIATED WITH RECEPTOR ACTIVATION  
CC (BY SIMILARITY).  
CC CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 570 570 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC SEQUENCE 576 AA; 65506 MW; 97D6A46CADCF42F CRC64;  
Query Match 35.7%; Score 943; DB 1; Length 576;  
Best Local Similarity 36.9%; Pred. No. 2, 6e-74;  
Matches 197; Conservative 102; Mismatches 189; Indels 46; Gaps 7;  
Qy 3 PMLAALALLPVSOGPEKRLNALLANYNTLRPVAEPEVRFGTLTQIIDV 62  
Db 25 PLCLLVLLVLLCETVQANPDAKLYDLLSNVRLRFPVSNNTDTVLVGLGRLSQIDL 84  
Qy 63 DEKNQLITNWLSEWDSYGVGVKDLRTPNKLWKPDLVLMVNSADEGDTG 122  
Db 85 NLKQDILITNWLSEWDSYGVGVKDLRTPNKLWKPDLVLMVNSADEGDTG 144  
Qy 123 YQNVVVRSGGCLYVPPGIFKSTCKMDIAWFFDDQCHDKMGFTYDGNQLDLV ---- 178  
Db 145 TMTKAILHYTGKVVVTPPAIFKSSCEIDRVYFPFDDQTCFMKFGSWYDGDIDLKHSQ 204  
Qy 179 -----LKDEAGGDLSDFITNGEWLQWPKKNTITVACCPYVVTFTIMIRRLIY 233

Db 205 KNDKXNKVEIGIDLREYVPSVEMDILGVPAERHEKYFCCAEYPYDIFENITLRKTLFY 264  
 Qy 234 FFNLIIVCVLISSMALLIGFTLPPSGEKLITGVITLISLVFLNLVETLPQVSDAIPLL 293  
 Db 265 TVNLIIPCGLISLVSVLYFYPADSGEKIALCISLLSQTNFFLLISEIIPSTSLALPLL 324  
 Qy 294 GTYENCIMFVASSVWLTVVVNLVHHRTADITHMPQWIKSVFLQWLWILRMSRPGKKIT 353  
 Db 325 GYKLLFTMLVGLSVVITIIILNHYKPSHKKWPIRSEFFIKLPLKLLMRVP----- 379  
 Qy 354 RKTIMMTRMELEKERSKSLANVLIDDDFRHGPPPPNS-TASTGNLGP-GCS-IF 410  
 Db 380 -KOLLRLAANKYGLKFKTKFQALMDQMNSGSGSDSLRRMQGRVGAGCGNGMH 438  
 Qy 411 RTDFRRFVRPSTMEDVGGGLGS-----HHREHLHLIRE 444  
 Db 439 VTTATNRF--SGLVGLGGGLSTLSGNGLPVSLVGLDDSLSDVAARKKYPFELEKAHN 496  
 Qy 445 LQFITARMKKADEAEELISOWKFAAMVVDRFCLFVFTLETTIATVAVLLSAPHI 498  
 Db 497 VMFIQHMHQDFNAEDQDWGFVAMVMDRLFLFLWFMIASLVGTFVILGEAPSL 550  
 RESULT 15  
 ACH6 HUMAN  
 ID ACH6 HUMAN STANDARD; PRT; 494 AA.  
 AC Q15825;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, last sequence update)  
 DT 10-OCT-2003 (Rel. 42, last annotation update)  
 DE Neuronal acetylcholine receptor protein, alpha-6 chain precursor.  
 GN CHRNA6.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Substantia nigra;  
 RX MEDLINE=97062879; PubMed=8906617;  
 RA Elliott K.J., Ellis S.B., Berckhan K.J., Urrutia A.,  
 RA Chavez-Noriega L.E., Johnson E.C., Velicelebi G., Harpold M.M.;  
 RT "Comparative structure of human neuronal alpha 2-alpha 7 and beta  
 RT 2-beta 4 nicotinic acetylcholine receptor subunits and functional  
 RT expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and  
 RT beta 4 subunits."; J. Mol. Neurosci. 7:217-228(1996).  
 RL J. Mol. Neurosci. 7:217-228(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC Groot Kormelink P.J.;  
 RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC Ebihara M., Ohba H., Yoshikawa T.;  
 RA "Alu and other elements in the promoter of human nAChR A6 gene  
 RT (CHNRA6) direct transcriptional repression."; Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=23388257; PubMed=12477932;  
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an  
 CC extensive change in conformation that affects all subunits and  
 CC leads to opening of an ion-conducting channel across the plasma  
 CC membrane.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U62435; AAB40113.1; -;  
 DR EMBL; Y16282; CAA76155.1; -;  
 DR EMBL; AB079251; BAC06855.1; -;  
 DR EMBL; AB079246; BAC06855.1; JOINED.  
 DR EMBL; AB079247; BAC06855.1; JOINED.  
 DR EMBL; AB079248; BAC06855.1; JOINED.  
 DR EMBL; AB079249; BAC06855.1; JOINED.  
 DR EMBL; AB079250; BAC06855.1; JOINED.  
 DR EMBL; BC014456; AAH14456.1; -;  
 DR EMBL; BC014456; AAH14456.1; -;  
 DR GenBank; HGNC:15963; CHRNA6.  
 DR MIM; 606888; -;  
 DR GO; GO:0005892; C:nicotinic acetylcholine-gated receptor-chan. . . ; TAS.  
 DR GO; GO:0015464; F:acetylcholine receptor activity; TAS.  
 DR GO; GO:0004889; F:nicotinic acetylcholine-activated cation-se. . . ; TAS.  
 DR GO; GO:0007165; P:signal transduction; TAS.  
 DR GO; GO:0007268; P:synaptic transmission; TAS.  
 DR GO; GO:0006810; P:transport; TAS.  
 DR InterPro; IPR006029; Neu\_chan\_memb.  
 DR InterPro; IPR006202; Neu\_chan\_LBD.  
 DR InterPro; IPR006201; Neu\_chan.  
 DR Pfam; PF02931; Neur\_chan\_LBD; 1.  
 DR Pfam; PF02932; Neur\_chan\_memb; 1.  
 DR PRINTS; PR00252; NRIONCHANNEL.  
 DR TIGRFAMs; TIGR00860; LIC; 1.  
 DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.  
 DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
 KW Transmembrane; Multigene family.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 494 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,  
 FT DOMAIN 26 239 ALPHA-6 CHAIN.  
 FT TRANSMEM 240 264 EXTRACELLULAR.  
 FT TRANSMEM 272 290 POTENTIAL.  
 FT TRANSMEM 306 327 POTENTIAL.  
 FT DOMAIN 328 465 CYTOPLASMIC.  
 FT TRANSMEM 466 484 POTENTIAL.  
 FT DISULFID 158 172 BY SIMILARITY.  
 FT DISULFID 222 223 ASSOCIATED WITH RECEPTOR ACTIVATION  
 FT CARBOHYD 54 54 (BY SIMILARITY).  
 FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 494 AA; 56898 MW; 1A437E6DE02ABFE CRC64;  
 Query Match 35.3%; Score 933; DB 1; Length 494;  
 Best Local Similarity 40.2%; Pred. No. 1.6e-73;  
 Matches 195; Conservative 83; Mismatches 159; Indels 48; Gaps 9;  
 Qy 23 EKRLNALLNANTLPRPVANSEPLEVFLGLTLOQIIVDEKNOLLITNWLSEWMDY 82  
 Db 34 EERLFKLSHYNQFIRPVENVSDPVTVHFEVAITQLANDEVNQLMNLWLRHWNDY 93

```

QY 83 NLEWDSYGVKDLRITPNKLPVLMYNSADEGPDGTQTNVVRSGSCLYVPPGI 142
Db 94 KLRWDPMEYDGIETLRVPAOKIWKPOILVYNNVAVGDFQVEGKTALLKYNMGITWTPPAI 153
QY 143 FKSTCKMDIAWFFPDQOHCMDKFGSWTYDGNQDLVLKDEAGG--DLSDFITNGEWYLIIG 200
Db 154 FKSSCPMDITFFPDHQCNSLXKFGSWTYDKAEIDLLI---IGSKVDNDFWENSEWEIID 210
QY 201 MPKKNITTYACCEPYVDVTFITMIRRTLYYFFNLIVPCVLISSMALLGFTLPPDSGE 260
Db 211 ASGYKHDIKYNCCEBIYTDITYSYIRRLPMFYTNLIIPCLFISFTLVLFVLPSCGE 270
QY 261 KLTILGVTLISLTVFLNLVAETLPQVSDAIPLLGTYENCIMFMVYASSVVLTVVVLNYHHR 320
Db 271 KVTLCISVLSLTVFLNLVAETLPQVSDAIPLLGTYENCIMFMVYASSVVLTVVVLNYHHR 330
QY 321 TADIHENPOWIKSVFQWLPWILMSRPGKKITRKTIMMTRMRELELKERSKSLIANV 380
Db 331 TPTHTMFRWKTVFLKLLPQVLLMRWPLDK-TRGT---GSDAVPRGLARRPAKGLAS- 385
QY 381 LDIDDDFRHGGP-----PPNSTASTGNLPGCSIPTDFRRSFRVPS--TMEDVG 428
Db 386 -----HGEPRHLKCFCHKSNELA-----TSKRRLSHQPLQWVVEN-- 422
QY 429 GGLGSHHRELHLILRELQFITARMKKADEEAELISDMKFAAMVYDRFCLFVFTLTFTIAT 488
Db 423 ---SEHSPEVEDVINSQVFAENWKSNETKEVEDDMKYAMVVDVFLWVFLWVFIIVCVFGT 479
QY 489 VAVLL 493
Db 480 AGLFL 484

```

Search completed: May 7, 2004, 11:36:33  
Job time : 10.3565 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:30:03 ; Search time 31.472 Seconds  
(without alignments)

5022.709 Million cell updates/sec

Title: US-09-303-232-6

Perfect score: 2640

Sequence: 1 MAPMLAALALLPVSEQ.....LFTIATVALLSAPHIIQ 501

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvrius:\*

16: sp\_bacteriapi:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2640	100.0	501	5 Q9XZ14	Q9xzi4 heliothis v
2	2194	83.1	494	5 Q8T7S1	Q8t7s1 drosophila
3	2181	82.6	494	5 Q8T7S2	Q8t7s2 drosophila
4	2177	82.5	494	5 Q8T7S3	Q8t7s3 drosophila
5	2177	82.5	494	5 Q86MN8	Q86mn8 drosophila
6	2176.5	82.4	509	5 Q8T7S0	Q8t7s0 drosophila
7	2156.5	81.7	523	5 Q8T7R9	Q8t7r9 drosophila
8	1833	69.4	554	5 Q9VL79	Q9vl79 drosophila
9	1803.5	68.3	496	5 Q9XZ13	Q9xzi3 heliothis v
10	1801.5	68.2	542	5 Q86MN7	Q86mn7 drosophila
11	1800.5	68.2	807	5 Q8T7V5	Q8t7v5 drosophila
12	1786.5	67.7	545	5 Q9VW19	Q9vwi9 drosophila
13	1705.5	64.6	525	5 Q81PE2	Q81pe2 drosophila
14	1258.5	47.7	502	11 Q9UHD6	Q9jhd6 mus musculus
15	1254.5	47.5	502	6 Q866A2	Q866a2 macaca mula
16	1245	47.2	555	13 Q7T2U0	Q7t2u0 fugu rubrip

17	1226.5	46.5	511	13 Q03481	Q03481 gallus gall
18	1222	46.3	509	13 Q800C7	Q800c7 brachydanio
19	1214	46.0	513	13 Q7T2R9	Q7t2r9 fugu rubrip
20	1211.5	45.9	486	13 Q7T2S0	Q7t2s0 fugu rubrip
21	1187.5	45.0	554	13 Q7T2T9	Q7t2t9 fugu rubrip
22	1145.5	43.4	480	5 Q81932	Q81932 caenorhabdi
23	1132	42.9	461	5 P91197	P91197 caenorhabdi
24	1131.5	42.9	474	13 Q7T2U1	Q7t2u1 fugu rubrip
25	1101	41.7	570	5 Q9VJ79	Q9vj79 drosophila
26	1085	41.1	554	5 Q62083	Q62083 caenorhabdi
27	1083.5	41.0	542	5 Q18556	Q18556 caenorhabdi
28	993.5	37.6	505	4 Q86U77	Q86u77 homo sapien
29	985	37.3	335	5 Q9NKD1	Q9nkd1 drosophila
30	982	37.2	537	5 Q8MUR0	Q8mur0 apis mellif
31	978	37.0	515	5 Q46133	Q46133 locusta mig
32	973.5	36.9	537	5 Q9U941	Q9u941 myzus persi
33	970.5	36.8	499	11 Q8R4G9	Q8r4g9 mus musculu
34	970.5	36.8	504	11 Q8BV44	Q8bv44 mus musculu
35	966.5	36.6	499	11 Q8VHH6	Q8vhh6 mus musculu
36	961	36.4	531	5 Q96632	Q96632 heliothis v
37	959.5	36.3	523	5 Q46128	Q46128 heliothis v
38	953.5	36.1	536	5 Q8T0Y9	Q8t0y9 aplysia cal
39	953	36.1	529	13 Q7T2S4	Q7t2s4 fugu rubrip
40	951.5	36.0	552	5 P91765	P91765 myzus persi
41	951.5	36.0	568	5 Q9NFR5	Q9nfr5 drosophila
42	946.5	35.9	545	5 Q96631	Q96631 heliothis v
43	941	35.6	532	5 Q9U940	Q9u940 myzus persi
44	939.5	35.6	536	5 Q8T9S0	Q8t9s0 aplysia cal
45	939	35.6	533	5 Q8WRS1	Q8wrs1 chilo suppr

## ALIGNMENTS

RESULT 1

Q9XZ14					
ID	Q9XZ14	PRELIMINARY;	PRT;	501 AA.	
AC	Q9XZ14;				
DT	01-NOV-1999 (TREMBLrel. 12, Created)				
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)				
DE	01-JUN-2003 (TREMBLrel. 24, Last annotation update)				
DE	Putative nicotinic acetylcholine receptor alpha 7-2 subunit.				
OS	Heliothis virescens (Noctuid moth) (Owlet moth).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;				
OC	Noctuidae; Heliothinae; Heliothis.				
OX	NCBI_TaxID=7102;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Schulte T., Oellers N., Adamczewski M.;				
RT	Putative alpha subunits of insect nicotinic acetylcholine receptors				
RT	more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than				
RT	to other insect nicotinic acetylcholine receptor alpha subunits.;				
RL	Submitted (APR-1999) to the EMBL/GenBank/DDBI databases.				
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).				
CC	-!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.				
DR	EMBL; AF143847; AAD32698.1; --				
DR	GO; GO:0016021; C:integral to membrane; IEA.				
DR	GO; GO:0005230; F:extracellular ligand-gated ion channel acti...; IEA.				
DR	GO; GO:0005216; F:ion channel activity; IEA.				
DR	GO; GO:0030594; F:neurotransmitter receptor activity; IEA.				
DR	GO; GO:0006811; P:ion transport; IEA.				
DR	GO; GO:0007269; P:synaptic transmission; IEA.				
DR	InterPro; IPR006201; Neur_chan				
DR	InterPro; IPR006029; Neur_chan_LBD				
DR	Pfam; PF02931; Neur_chan_LBD; 1.				
DR	Pfam; PF02932; Neur_chan_memb; 1.				
DR	PRINTS; PR00252; NRIONCHANNEL.				
DR	TIGRFAMs; TIGR00860; LIC; 1.				
DR	PROSITE; PS00236; NEUROTRF ION CHANNEL; 1.				
KW	Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;				
KW	Transmembrane.				

SQ SEQUENCE 501 AA; 56704 MW; 43CB0DC3960C78AB CRC64;  
Query Match 100.0%; Score 2640; DB 5; Length 501;  
Best Local Similarity 100.0%; Pred. No. 6.9e-242;  
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAPMLAALALLPVSEQGPHEKRLNALLANNVTLEPRVANESEPLEVRFGTLTQQII 60  
DB 1 MAPMLAALALLPVSEQGPHEKRLNALLANNVTLEPRVANESEPLEVRFGTLTQQII 60  
QY 61 DVDEKNQLLITNWLSEWYNDYLRWDSYGGVVKDLRITPNKLMKPDVLMYNSADSGFD 120  
DB 61 DVDEKNQLLITNWLSEWYNDYLRWDSYGGVVKDLRITPNKLMKPDVLMYNSADSGFD 120  
QY 121 GTYTQNVVVRSGGSLVPPGIFKSTCKMDIAWPPDDQCHDMKFGSWTYDGNQDLVLK 180  
DB 121 GTYTQNVVVRSGGSLVPPGIFKSTCKMDIAWPPDDQCHDMKFGSWTYDGNQDLVLK 180  
QY 181 DEAGGDLSDFTNGEWYLGMPGKNTITVACCEPVVDVTFITMIRRTLYYFENLIVP 240  
DB 181 DEAGGDLSDFTNGEWYLGMPGKNTITVACCEPVVDVTFITMIRRTLYYFENLIVP 240  
QY 241 CVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGTYNFI 300  
DB 241 CVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGTYNFI 300  
QY 301 MFWVASSVLTVVVLYNHHTADITHMPQWIKSVFLOLWFLWILMSRPGKIKTRKTMN 360  
DB 301 MFWVASSVLTVVVLYNHHTADITHMPQWIKSVFLOLWFLWILMSRPGKIKTRKTMN 360  
QY 361 TRMRELEKERSKSLANVLIDDDFRHGPPPNSTASTGNLPGGCSIFRTDFRSFVR 420  
DB 361 TRMRELEKERSKSLANVLIDDDFRHGPPPNSTASTGNLPGGCSIFRTDFRSFVR 420  
QY 421 PSTMEDVGGGLSHHRELHLILRELQFITARMKKADEAEELISDWKFAAMVVDRCFLVF 480  
DB 421 PSTMEDVGGGLSHHRELHLILRELQFITARMKKADEAEELISDWKFAAMVVDRCFLVF 480  
QY 481 TLFTIIATVALLSAPHIIQV 501  
DB 481 TLFTIIATVALLSAPHIIQV 501

## RESULT 2

Q8T7S1 ID Q8T7S1 PRELIMINARY; PRT; 494 AA.  
AC Q8T7S1;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type III.  
DE NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.  
GN Drosophila melanogaster (Fruit fly).  
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
MEDLINE=21969411; PubMed=11973307;  
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;  
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,  
RT Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a  
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-  
RT Mediated A-to-I Pre-mRNA Editing.";  
RL Genetics 160:1519-1533(2002).  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
DR EMBL; AF321447; AAM13394.1; -;  
DR FlyBase; FBgn0032151; nacr-alpha-30D.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.

DR GO; GO:0005216; F:ion channel activity; IEA.  
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.  
DR GO; GO:0006811; P:ion transport; IEA.  
DR GO; GO:0007268; P:synaptic transmission; IEA.  
DR InterPro; IPR006201; Neur\_chan channel.  
DR InterPro; IPR006202; Neur\_chan\_LBD.  
DR InterPro; IPR006029; Neur\_chan\_LBD; 1\_memb.  
DR Pfam; PF02931; Neur\_chan\_LBD; 1\_memb.  
DR Pfam; PF02932; Neur\_chan\_memb; 1.  
DR PRINTS; PR00252; NRIONCHANNEL.  
DR TIGRFAMS; TIGR00860; LIC; 1.  
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.  
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;  
KW Transmembrane.  
SQ SEQUENCE 494 AA; 56113 MW; 48327537229573FF CRC64;

Query Match 83.1%; Score 2194; DB 5; Length 494;  
Best Local Similarity 82.6%; Pred. No. 1.7e-199;  
Matches 416; Conservative 31; Mismatches 37; Indels 20; Gaps 3;

QY 2 APMLAALAL-----LALLPVSEQGPHEKRLNALLANNVTLEPRVANESEPLEVRFGLT 55  
DB 3 SPUPASUSLVLLIFLAIKESQGPHEKRLNALLANNVTLEPRVANESEPLEVRFGLT 62  
QY 56 LQQTIDVDEKNQLLITNWLSEWYNDYLRWDSYGGVVKDLRITPNKLMKPDVLMYNSA 115  
DB 63 LQQTIDVDEKNQLLITNWLSEWYNDYLRWDSYGGVVKDLRITPNKLMKPDVLMYNSA 122  
QY 116 DEFGDGYQNVVVRSGGSLVPPGIFKSTCKMDIAWPPDDQCHDMKFGSWTYDGNQL 175  
DB 123 DEFGDGYHTNIVVVKHSGSLVPPGIFKSTCKMDITWFPDDQCHDMKFGSWTYDGNQL 182  
QY 176 DLVLKAGGDLSDFTNGEWYLGMPGKNTITVACCEPVVDVTFITMIRRTLYYF 235  
DB 183 DLVNSDGGDLSDFTNGEWYLGMPGKNTITVACCEPVVDVTFITMIRRTLYYF 242  
QY 236 NLIIVPCVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGT 295  
DB 243 NLIIVPCVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGT 302  
QY 296 YFCIMFWASSVLTVVVLYNHHTADITHMPQWIKSVFLOLWFLWILMSRPGKIKTRK 355  
DB 303 YFCIMFWASSVLTVVVLYNHHTADITHMPQWIKSVFLOLWFLWILMSRPGKIKTRK 362  
QY 356 TIMMTRMRELEKERSKSLANVLIDDDFRHGPPPNSTASTGNLPGGCSIFRTDFR 415  
DB 363 TILLSNRKLEKERSKSLANVLIDDDFRH-----TISGQTAIGSS----- 408  
QY 416 RSFVRPSTMEDVGGGLSHHRELHLILRELQFITARMKKADEAEELISDWKFAAMVVD 475  
DB 409 ASFGRTTVEHHTAIGNHKDLHLILRELQFITARMKKADEAEELISDWKFAAMVVD 468  
QY 476 CLFVFTLTIIATVALLSAPHIIQV 501  
DB 469 CLFVFTLTIIATVALLSAPHIIQV 494

## RESULT 3

Q8T7S2 ID Q8T7S2 PRELIMINARY; PRT; 494 AA.  
AC Q8T7S2;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type II.  
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.



```

QY 416 RSFVRPSTMEDVGGGLGSHHREHLILRELQFITARMKKADEAEELISDWKFAAMVVDRF 475
Db 409 ASFGPRTTVEHHTAICGNHKLHLILRELQFITARMKKADDEAEELIGDWKFAAMVVDRF 468
QY 476 CLFVFTLFTIIATVAVLLSAPHIIQV 501
Db 469 CLIVFTLFTIIATVAVLLSAPHIIQV 494

RESULT 5
Q86MN8 Q86MN8 PRELIMINARY; PRT; 494 AA.
AC Q86MN8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nicotinic acetylcholine receptor subunit Dalpha6 precursor.
GN NACRALPHA-30D.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Millar N.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ554209; CAD86935.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
DR SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 494 NICOTINIC ACETYLCHOLINE RECEPTOR SUBUNIT
FT DALPHA6.
FT SEQUENCE 494 AA; 55980 MW; 4DFC572139587070 CRC64;

Query Match 82.5%; Score 2177; DB 5; Length 494;
Best Local Similarity 81.6%; Pred. No. 6.8e-198;
Matches 413; Conservative 35; Mismatches 38; Indels 20; Gaps 3;

QY 2 APMLAALAL-----LALLPVSEOGPEKRLINALLANYNTLRPVANESPLEVRGLT 55
Db 3 SPLPASUSLVLLIFLAIKESCGPHEKRLINLLSTYNTLRPVANESPLEVRKGLT 62
QY 56 LQIIVDEKNQILITNWLSEWNDYLNLRNDSEYGVKDLRITPNKMKPDLVMTNSA 115
Db 63 LQIIVDEKNQILITNWLSEWNDYLNLRNWEYGVKDLRITPNKMKPDLVMTNSA 122
QY 116 DEGPDTGYQTNVVRSGSCLVYPPGIFKSTCKMDIANWFPDDOHCMDKFGSWYDGNL 175
Db 123 DEGPDTGYHTSVVVKHGSCCLVYPPGIFKSTCKMDITWFPDDOHCMDKFGSWYDGNL 182
QY 176 DLVLKDGAGDLSDFITNGEWYLMGPKKNTITYACCPYVDVTFITMIRRTLYYFF 235
Db 183 DLVLSSDEGGDLSDFITNGEWYLMGPKKNTITYACCPYVDVTFITQIRRTLYYFF 242
QY 236 NLIIVPCVLISMAILGFTLPDSGEKLTGLVITLLSLTVFLNVAETLPQVSDAIPLLGT 295
Db 243 NLIIVPCVLISMAILGFTLPDSGEKLTGLVITLLSLTVFLNVAESNFTSDAVPLIGT 302
QY 296 YFCIMFVWASSVVLTVVVLNYHRTADIHMPQWIKSVFLQWLFWILMRSPGKKITRK 355

```

```

Db 303 YFCIMFVWASSVVLTVVVLNYHRTADIHMPQWIKSVFLQWLFWILMRSPGKKITRK 362
QY 356 TIMMTRMRELEKERSKSLANVLIDDDDRHGPPPPNSTAGNLGPGCSIFRTDPR 415
Db 363 TILLSNRMRLEKERSKSLANVLIDDDDRH-----TISGSTAIGSS----- 408
QY 416 RSFVRPSTMEDVGGGLGSHHREHLILRELQFITARMKKADEAEELISDWKFAAMVVDRF 475
Db 409 ASFGPRTTVEHHTAICGNHKLHLILRELQFITARMKKADDEAEELIGDWKFAAMVVDRF 468
QY 476 CLFVFTLFTIIATVAVLLSAPHIIQV 501
Db 469 CLIVFTLFTIIATVAVLLSAPHIIQV 494

RESULT 6
Q8T7S0 Q8T7S0 PRELIMINARY; PRT; 509 AA.
AC Q8T7S0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type IV.
GN NACRALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes
RT Dalpha6, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.
RL Genetics 160:1519-1533(2002).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321448; AA013395.1; -.
DR FlyBase; FBgn032151; nACR-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_memb.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
DR Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 509 AA; 57987 MW; BE8D8E0198E0C2BD CRC64;

Query Match 82.4%; Score 2176.5; DB 5; Length 509;
Best Local Similarity 80.2%; Pred. No. 8e-198;
Matches 418; Conservative 31; Mismatches 37; Indels 35; Gaps 4;

QY 2 APMLAALAL-----LALLPVSEOGPEKRLINALLANYNTLRPVANESPLEVRGLT 55
Db 3 SPLPASUSLVLLIFLAIKESCGPHEKRLINLLSTYNTLRPVANESPLEVRKGLT 62
QY 56 LQIIVDEKNQILITNWLSEWNDYLNLRNDSEYGVKDLRITPNKMKPDLVMTNSA 100
Db 63 LQIIVDEKNQILITNWLSEWNDYLNLRNWEYGVKDLRITPNKMKPDLVMTNSA 122
QY 101 PNKLWKPDLVMTNSADEGFDGTQYTNVVRSGSCLVYPPGIFKSTCKMDIANWFPDDOHC 160

```

```

Db 123 PNKLKEDVLNYSADSGFGTHTNIVVRSGLVYVPGIFKSTCKMDITWFPDDQH 182
QY 161 CDWKFGSWTDGNQDLVLKDEAGDLSDFITNGEWYILGMPGKNTITYACCPYVDV 220
Db 183 CEMKFGSWTDGNQDLVLNSEDGDLSDFITNGEWYLLAMPKNTIVVACCPYVDI 242
QY 221 TPTIMRRRTLYFFENLIVPCVLISSMALLGFTLPDPSGKILGVITLISLVFLNVA 280
Db 243 TPTIQRRRTLYFFENLIVPCVLISSMALLGFTLPDPSGKILGVITLISLVFLNVA 302
QY 281 ETLFQVSDAIPLLGTGYNCIMFVASSVLTVVVLYNHHRTADITHEMPOMKSVFLQWLP 340
Db 303 ETLFQVSDAIPLLGTGYNCIMFVASSVLTVVVLYNHHRTADITHEMPOMKSVFLQWLP 362
QY 341 WILMRSPGKTKTRKTIIMNTRMRELEKERSKSLIANVLDDDDPRHGGPPPNSTAST 400
Db 363 WILMRSPGKTKTRKTIIMNTRMRELEKERSKSLIANVLDDDDPRH-----TISG 415
QY 401 GNLPGGCSIFRTDPRRSFVRPSTMEDVGGIGSHHRELHLILFELQFITARMKKADEAE 460
Db 416 SQTAGSS-----ASFGRTTVEEHHTAIGNHKLHLILFELQFITARMKKADEAE 468
QY 461 LISDWKFAAMVDRFCLFVFTFTTIATVAVLLSAPHIIVQ 501
Db 469 LIGDWKFAAMVDRFCLFVFTFTTIATVAVLLSAPHIIVQ 509

RESULT 7
Q8T7R9 PRELIMINARY; PRT; 523 AA.
AC Q8T7R9
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DE Nicotinic acetylcholine receptor Dalphae subunit variant type V.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culiceto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalphae, Dalphae and Dalphae7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533 (2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY) .
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321449; AAM1396.1; -.
DR FlyBase; FBgn0032151; nACR-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:000594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006202; Neur_chan.
DR InterPro; IPR006029; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTRP_ION_CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 523 AA; 59110 MW; 1C200AF74F87F841 CRC64;

```

```

Query Match 81.7%; Score 2156.5; DB 5; Length 523;
Best local Similarity 77.6%; Pred. No. 6.6e-196;
Matches 415; Conservative 32; Mismatches 39; Indels 49; Gaps 4;

QY 2 APLMALAL-----LALLPVSQGHKEKELLNALLANYNTLRPPVANESEPLEVRFGLT 55
Db 3 SPIPASLSLVLLIFALIKESCGPHEKELLNLLSTYNTLRPPVANESEPLEVRFGLT 62
QY 56 LQOIIDVDEKNQLLITNINLSLEWNDYNLEWNSYGGVKKDLRITPNKLPDVLNYS 115
Db 63 LQOIIDVDEKNQLLITNINLSLEWNDYNLEWNSYGGVKKDLRITPNKLPDVLNYS 122
QY 116 DEFGDGTQYNNVVRSGSCLYVPPGIFKSTCKMDIAMFPDQCHCKMFGSWTDGNQL 175
Db 123 DEFGDGTQYNNVVRSGSCLYVPPGIFKSTCKMDIAMFPDQCHCKMFGSWTDGNQL 182
QY 176 DLVLKDEAGDLSDFITNGEWYILGMPGKNTITYACCPYVDVFTIMRRRTLYFF 235
Db 183 DLVLNSEDGDLSDFITNGEWYLLAMPKNTIVVACCPYVDITFTIQIRRTLYFF 242
QY 236 NLIVPCVLISSMALLGFTLPDPSGKILT-----GV 266
Db 243 NLIVPCVLISSMALLGFTLPDPSGKILT-----GV 302
QY 267 TILLSITVFLNVAETLPQVSDAIPLLGTGYNCIMFVASSVLTVVVLYNHHRTADIHE 326
Db 303 TILLSITVFLNVAETLPQVSDAIPLLGTGYNCIMFVASSVLTVVVLYNHHRTADIHE 362
QY 327 MPQWIKSVFLOLFWILRMSRPGKTKTRKTIIMNTRMRELEKERSKSLIANVLDDDD 386
Db 363 MPWIKSVFLOLFWILRMSRPGKTKTRKTIIMNTRMRELEKERSKSLIANVLDDDD 422
QY 387 FRHGPPPNSTASTGNLPGFCSIFRTDPRRSFVRPSTMEDVGGIGSHHRELHLILRELQ 446
Db 423 FRH-----TISGSAIGSS-----ASFGRTTVEEHHTAIGNHKLHLILRELQ 468
QY 447 FITARMKKADEAEABLISDWKFAAMVDRFCLFVFTFTTIATVAVLLSAPHIIVQ 501
Db 469 FITARMKKADEAEABLISDWKFAAMVDRFCLFVFTFTTIATVAVLLSAPHIIVQ 523

RESULT 8
Q9VL79 PRELIMINARY; PRT; 554 AA.
AC Q9VL79
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE CG4128 protein.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Gelink S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

```

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez T.J., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lin X.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Mobarri C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*;"  
RL Science 287:2185-2195(2000).  
RN [2]  
RN SEQUENCE FROM N.A.  
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banzon J., An H., Baldwin D., Bantz J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Chapple M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattel B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
RT "sequencing of *Drosophila melanogaster* genome;"  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE FROM N.A.  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kanink J.S., Prochnik S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
RT "Annotation of *Drosophila melanogaster* genome;"  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RN SEQUENCE FROM N.A.  
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RN SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AEO03626; AAF52817.2; -;  
DR FlyBase; FBgn0032151; nAcr-alpha-30D.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.  
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.  
DR GO; GO:0006811; P:ion transport; IEA.  
DR InterPro; IPR006201; Neur\_channel.  
DR InterPro; IPR006202; Neur\_chan\_LBD.  
DR Pfam; PF02931; Neur\_chan\_LBD.  
DR Pfam; PF02932; Neur\_chan\_LBD; 1.  
DR PRINTS; PR00252; NRIONCHANNEL.  
DR TIGRFAMs; TIGR00860; LIC; 1.

DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.  
SQ SEQUENCE 554 AA; 63236 MW; D6C63EB7F5B32A35 CRC64;  
Query Match 69.4%; Score 1833; DB 5; Length 554;  
Best Local Similarity 65.0%; Pred. No. 4e-165;  
Matches 370; Conservative 36; Mismatches 67; Indels 96; Gaps 8;  
QY 4 MLAALALLPVSEOGPEKRLINALLANTLERPVANESBPFLVRFGLTLQQLIIVD 63  
DB 11 LFVLLFLAIKESQCPHEKRLINHLSTYNTLERPVANESBPFLVRFGLTLQQLIIVD 70  
QY 64 EKQQLITNLSLEWNYLNWDSRYGVGKDLRITPNKLPKPDVLMYNSADEGFGTY 123  
DB 71 EKQQLITNLSLEWNYLNWDSRYGVGKDLRITPNKLPKPDVLMYNSADEGFGTY 130  
QY 124 QTNVVRSGGCLYVPGPIKSKMDIAWFPDDQDCHDMKFGSWTYDGNQLDLVLKDEA 183  
DB 131 HTNIVKHNKSGCLYVPGPIKSKMDITWFPDDQDCHDMKFGSWTYDGNQLDLVLNSED 190  
QY 184 GGLSDPITNGEWYLGIM-PGKNTITYACCP-----EPYVDVTFITMIRRR 229  
DB 191 GGLSDPITNGEWYLGIM-PGKNTITYACCP-----EPYVDVTFITMIRRR 239  
QY 230 TLYFFNLIIVPCVLISSMALL-----GFTLPDSEKLT----- 263  
DB 240 ILFFQFNCAACANLIDGPTGLHIAAGFRETDAGRNYTTIINSISKEPCRRVHADNVGCCS 299  
QY 264 -----LG-----VTLLSLTVLNLVAETLPQVSDAIP 292  
DB 300 SYRYTHILKRYSLKYGPIDRTDHYQCRARTVLLSLTVLNLVAETLPQVSDAIP 359  
QY 293 LGTYFNCIMFWASSVVLTVVLNHYHRTADIHEMPQWIKSVFLQWLPWILMRSPGKI 352  
DB 360 LGTYFNCIMFWASSVVLTVVLNHYHRTADIHEMPQWIKSVFLQWLPWILMRSPGKI 419  
QY 353 TRKTMNTMRLELKERSSKSLANVLIDDDFRGPPPPNSTAGTGNLGPCCSIFRT 412  
DB 420 TRKTMNTMRLELKERSSKSLANVLIDDDFRH-----TISGQTAIGSS----- 468  
QY 413 DFRSFVRPSTMDVGGGLSHHRELHLILRLQFITARMKKADEAEELISDWKFAAMV 472  
DB 469 ---ASFGRPTVBEHHTAIGCNHKLHLILRLQFITARMKKADEAEELISDWKFAAMV 525  
QY 473 DRFCFLVFTLTIIATVALLSAPHIIIVQ 501  
DB 526 DRFCFLVFTLTIIATVALLSAPHIIIVQ 554

## RESULT 9

Q9XZ13  
ID Q9XZ13 PRELIMINARY; PRT; 496 AA.  
AC Q9XZ13;  
DT 01-NOV-1999 (TREMELrel. 12, Created)  
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
DE 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DE Putative nicotinic acetylcholine receptor alpha 7-1 subunit.  
OS Heliothis virescens (Noctuid moth) (Owlet moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;  
OC Noctuidae; Heliothinae; Heliothis.  
OX NCBI\_TaxID=7102;  
RN [1]  
RN SEQUENCE FROM N.A.  
RA Schulte T., Oellers N., Adamczewski M.;  
RT "Putative alpha subunits of insect nicotinic acetylcholine receptors  
RT more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than  
RT to other insect nicotinic acetylcholine receptor alpha subunits.";  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
DR EMBL; AF143846; AAD32697.1; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.









Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
"Sequencing of Drosophila melanogaster genome";  
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE FROM N.A.  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Bernan B., Carlson J.W., Celniker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
RA "Annotation of Drosophila melanogaster genome";  
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[4]  
SEQUENCE FROM N.A.  
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[5]  
SEQUENCE FROM N.A.  
RA FlyBase;  
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RA EMBL; AE003311; AAF48950.2; nAcR-alpha-18C.  
RA FlyBase; FBgn0031014; nAcR-alpha-18C.  
RA GO; GO:0016020; C:membrane; IEA.  
RA GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . . ; IEA.  
RA GO; GO:0030594; F:neurotransmitter receptor activity; IEA.  
RA GO; GO:0006811; P:ion transport; IEA.  
RA InterPro; IPR006201; Neur\_channel.  
RA InterPro; IPR006202; Neur\_chan\_LBD.  
RA InterPro; IPR006029; Neur\_channel\_memb.  
RA Pfam; PF02931; Neur\_chan\_LBD; 1.  
RA Pfam; PF02932; Neur\_chan\_memb; 1.  
RA PRINTS; PR00252; NRIONCHANNEL.  
RA TIGRFS; TIGR00860; LIC; 1.  
RA PROSITE; PS00236; NEUROTR ION CHANNEL; 1.  
RA SEQUENCE 545 AA; 61517 MW; 7B83025107A66209 CRC64;  
Query Match 67.7%; Score 1786.5; DB 5; Length 545;  
Best Local Similarity 66.7%; Pred. No. 1e-160;  
Matches 343; Conservative 56; Mismatches 70; Indels 45; Gaps 5;  
20 GPHEKLLNALLANYTLERPVANESPLEVREGLTQQIIDVDKXNQLITNWLISLEW 79  
39 GPHEKLLHALLDYNLSERPVVNSDPLQSLGTLMIQIIDVDKXNQLITNWLKLEW 98  
80 NDYNLRWNSSEYGVKDLITPNKLPDVLNYSADGFDGYQYNVVRSGSCLYVP 139  
99 NDYNLRWNSSEYGVKDLITPNKLPDVLNYSADGFDGYQYNVVRSGSCLYVP 158  
140 PGIFKSTKMDJAWFPDQDCMKFGSWTYDGNQLDLVKBAGDLSDFITNGEWYLI 199  
159 PGIFKSTKIDITWFPDQDCMKFGSWTYDGFQDLQDDBAGDLSDFITNGEWYLI 218  
200 GMPGKQNTIYACCPYVDVITIMRRITLYFFNLVPCVLISSMALLGFTLPDGS 259  
219 GYGVKRNIEYVNCPEYDITFALLIRKTLTYFFNLVPCVLISMAALLGFTLPDGS 278  
260 EKLTLGVITLLSLTVPLNVAETLPVDSDAIPLLGYFNCIMFWASSVVLTVVLYHH 319  
279 EKLTLGVITLLSLTVPLNVAETLPVDSDAIPLLGYFNCIMFWASSVVLTVVLYHH 338  
320 RTADIHMPQWTKSVFLOLWPLMRSPGK-----KIFKTIIMTRREL 368  
339 RNPDTHEMSEWRVFLYVLPCLLRQVGVYECPPPPSSSSSASGEKKQIQNVEL 398  
369 KERSKSLIANVLIDDDDRHGGPPPNSTAGNLGPGCSIFRTPFRFRVAPSTMED-- 426  
399 KERSKSLIANVLIDDDDR-----CNHRCASATLPHQTYTYTYRQG-----DQGS 446  
427 -----VGGL-----GSHREHLILRELQFTARMKKADEAEILISDWK 466  
447 VGFVGPAGVPVGRHLHEAHSHTCLTSSARYELAILKELRWITEQLKKEDETSITRDWK 506

Qy 467 FAAMVDRFCLEFVFTFTIATVAVLSAPHIIV 500  
Db 507 FAAMVDRCLIIFTFTIATVAVLSAPHFIV 540

## RESULT 13

Q8IPE2 PRELIMINARY; PRT; 525 AA.  
AC Q8IPE2  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CG4128-PC.  
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abrial J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Durbin K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spiers E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svetskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RA "The genome sequence of Drosophila melanogaster";  
RA Science 287:2185-2195(2000).  
[2]  
RP SEQUENCE FROM N.A.  
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,  
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez C., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Pacleeb J., Paragas V., Park S., Patel S., Pfeiffer B.,



```
Db 130 VLVNASGHCQYLPPIGFIKSSCYIDVRWFFEDVQCKLKFGSKSYGWSLDLQMQE--AD 186
Qy 187 LSFDTNGEWYLGMPGKNTITYACCPYVDVTFIMIRRTLYYFENLIVPCVLSS 246
Db 187 ISSIPNGEWDLGIPGRKNEKFYECCKEPYDVTYTMRRRTLYYGLNLLIPCVLISA 246
Qy 247 MALLGFTLPDPSGKLTGVTLLSLTVFLNLVAETLPQVSDAIFLLGTYFNCIMFWAS 306
Db 247 LALLVELPADSGEKISIGITVLLSLTVFLNLVAEIMPATSDSVPLIAQYFASMTIIVGL 306
Qy 307 SVLVTVVVLNHHHTADIHEMPQWIKSVFLOLWFLMRLMSPGKKITRKTIMMTRREL 366
Db 307 SVVTVIVLVYHHHDPDGGKMPKWTIRILLNWCWFLRMKRPGEKVRPAQCHKPRCSL 366
Qy 367 ELKERS--KSLANVLDIDDDFR-----HGPPPPNSTASTGNLPGGCSIFRTDFRSFV 417
Db 367 ASVELSAGAPPPSNGNLLYI--GFRGLEGMHCAPTDPDSGVVCGRL--ACSPTHDEHLMH 422
Qy 418 FVRPSTMEDVGGGLGSHHRELHLILRELQFITARMKXADEAEELISDWKFAAMVVD3FCL 477
Db 423 GTHPSDGP-----DLAKILEEVRYIANFRQDESEVICSEWKFACVVDRLCL 472
Qy 478 FVFTLFTIATVALLSAPHII 499
Db 473 MAFSVFTIICITIGILMSAPNFV 494
```

## RESULT 15

```
Q866A2 PRELIMINARY; PRT; 502 AA.
AC Q866A2;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Nicotinic acetylcholine receptor alpha7.
GN CHRNA7.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Proskocil B.J., Sekhon H.S., Keller J.A., Jia Y., Blakely R.D.,
RA Lindstrom J., Spindel E.R.;
RT "An Intrinsic Non-neuronal Nicotinic Cholinergic Signaling System in
RL Monkey Lung Airway Epithelium.";
DR EMBL; AF486623; AAC084497.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . ; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0008811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur channel.
DR InterPro; IPR006202; Neur chan LBD.
DR InterPro; IPR006029; Neu Channel LBD.
DR Pfam; PF02931; Neur chan_LBD; 1_memb.
DR Pfam; PF02932; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRfams; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Receptor.
SQ SEQUENCE 502 AA; 56429 MW; 213C8A28224AC4A CRC64;
```

```
Query Match 47.5%; Score 1254.5; DB 6; Length 502;
Best Local Similarity 49.6%; Pred. No. 3.4e-110;
Matches 249; Conservative 79; Mismatches 147; Indels 27; Gaps 7;
Qy 8 LALLA-LLPVSGPHEKELLNALLANNYTLRPERVANESEPLEVRPLTLOQIIVDEKN 66
Db 10 LALAASLLHVSLSQGBFQRLKYLKVNPNLPERPVANDSQPLTVYFSLQLQIMDVDEKN 69
```

```
Qy 67 QLLITNIWLSLEWNDYNIWRNDSEYGGVKDILRITPNKLWPKDVLMYNSADEGFGTYQTN 126
Db 70 QVLTATNIWQMSWTDHYLQMNVSSEYGVKTVRFDPDQIWKPDILLYNSADERFDATEHTN 129
Qy 127 VVVSRSGLVAVPGIEKSTCKMDIANWFFPDQCHDKMFGSWTYDGNOLDLVLKDEAGD 186
Db 130 VLVNSSHQCQLPPGIIFKSSCYIDVRWFFEDVQCKLKFGSKSYGWSLDLQMQE--AD 186
Qy 187 LSFDTNGEWYLGMPGKNTITYACCPYVDVTFIMIRRTLYYFENLIVPCVLSS 246
Db 187 ISSIPNGEWDLGIPGRKNEKFYECCKEPYDVTYTMRRRTLYYGLNLLIPCVLISA 246
Qy 247 MALLGFTLPDPSGKLTGVTLLSLTVFLNLVAETLPQVSDAIFLLGTYFNCIMFWAS 306
Db 247 LALLVELPADSGEKISIGITVLLSLTVFLNLVAEIMPATSDSVPLIAQYFASMTIIVGL 306
Qy 307 SVLVTVVVLNHHHTADIHEMPQWIKSVFLOLWFLMRLMSPGKKITRKTIMMTRREL 366
Db 307 SVVTVIVLVYHHHDPDGGKMPKWTIRILLNWCWFLRMKRPGEKVRPAQCHKPRCSL 366
Qy 367 ELKERS--KSLANVLDIDDDFR-----HGPPPPNSTASTGNLPGGCSIFRTDFRSFV 419
Db 367 ASVELSAGAPPPSNGNLLYI--GFRGLEGMHCAPTDPDSGVVCGRL--ACSPTHDEHLMH 413
Qy 420 RPTMEDV--GGGLGSHHRELHLILRELQFITARMKXADEAEELISDWKFAAMVVD3FCL 477
Db 414 -PTHDEHLLGGQPPEGDPDLAKILEEVRYIANFRQDESEAVCSWKFACVVDRLCL 472
Qy 478 FVFTLFTIATVALLSAPHII 499
Db 473 MAFSVFTIICITIGILMSAPNFV 494
```

Search completed: May 7, 2004, 11:38:39  
Job time : 32.472 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2004, 11:50:34 ; Search time 3498.21 Seconds  
(without alignments)  
4276.744 Million cell updates/sec

Title: US-09-303-232-6  
Perfect score: 2640  
Sequence: 1 MAPMLAALALLLPVSEQ.....LFTIATVAVLLSAPHIIVQ 501

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cpn2.1/USPTO\_spool/US09303232/runat\_07052004.101110.23893/app\_query.fasta\_1.2261  
-DB=EST -QPMI=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09303232 @CCN 1.1 4577 @runat\_07052004.101110.23893 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_estl.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rtd.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gssl.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C					
1	1273	48.2	885	12	BG632919 GH16126.3
2	970.5	36.8	1864	11	AK053497 Mus muscu
3	970.5	36.8	2916	11	AK051730 Mus muscu
4	970.5	36.8	3126	11	AK080415 Mus muscu
5	957.5	36.3	1436	29	AY402873 Homo sapi
6	944.5	35.8	1436	29	AY402875 Mus muscu
7	924	35.0	2010	11	AK080475 Mus muscu
8	922.5	34.9	2296	14	CD013901 90134548
9	914	34.6	1454	29	AY402878 Mus muscu
10	910.5	34.5	4037	11	AK049722 Mus muscu
11	910.5	34.5	4046	11	AK051742 Mus muscu
12	909	34.4	1201	9	AL530299
13	905	34.3	1374	29	AY406230
14	890.5	33.7	2940	11	AK034228 Mus muscu
15	890.5	33.7	3230	11	AK083157 Mus muscu
16	886.5	33.6	4290	11	AK029177 Mus muscu
17	886	33.6	1374	29	AY406232 Mus muscu
18	884	33.5	1442	29	AY402876 Mus muscu
19	882.5	33.4	1781	14	CD013889
20	881	33.4	1442	29	AY402877
21	865.5	32.8	1374	29	AY406231
22	853	32.3	4589	11	AK030464 Mus muscu
23	849	32.2	922	13	BU915857
24	847	32.1	607	9	AI292581
25	846.5	32.1	1436	29	AY402874
26	838.5	31.8	3483	11	AK081254 Mus muscu
27	791.5	30.0	1466	14	CD013891 90139069
28	776.5	29.4	2257	11	AK017571 Mus muscu
29	770	29.2	1751	14	CD013905 90134608
30	770	29.2	2110	14	CD013902
31	770	29.2	2110	14	CD013903
32	770	29.2	2513	11	AK033068
33	769.5	29.1	2074	11	AK087554
34	759	28.7	908	13	BU149265
35	748.5	28.4	1034	13	BU149265
36	737.5	27.9	1596	14	CD013888
37	716	27.1	1603	29	AY411327
38	703.5	26.6	833	14	CB245337
39	693	26.2	902	29	AY407184
40	691.5	26.2	2948	11	AK081344
41	686	26.0	902	29	AY407186
42	682.5	25.9	615	14	CB149460
43	679.5	25.7	1603	29	AY411329
44	672.5	25.5	1835	11	AK010496
45	666	25.2	1556	14	CD013895

# ALIGNMENTS

RESULT 1  
BG632919/c  
LOCUS  
DEFINITION  
885 bp mRNA linear EST 23-APR-2001  
GH16126.3prime GH Drosophila melanogaster head polyA Drosophila  
melanogaster cDNA clone GH16126.3 similar to CG4128: FBan0004128  
'ion channel', located on: 2L 30D1-30E1:: 04/10/2001, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
BG632919  
BG632919.1 GI:13758409  
EST.  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 885)  
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,  
Lewis, S. and Rubin, G.M.  
BDGP/HMMI Drosophila EST Project  
Unpublished (2001)  
Other ESTs: GH16126.5prime  
Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd. Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
Based upon the presence of a xhoI site followed by a run of 14 or  
more T residues at the beginning of the sequence, this clone was  
polyadenylated. The resulting poly-T sequence has been removed. hit  
genomic AE003511: arm:X [18792641,19136447]  
estimated-cyto:18A3-18C6: 04/10/2001  
Plate: GH.161 row: C column: 2  
High quality sequence stop: 784.  
Location/Qualifiers  
1. .885  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clones="GH16126"  
/sex="male and female"  
/dev\_stage="adult"  
/lab\_host="DH5 - alpha"  
/clone\_lib="GH Drosophila melanogaster head pOT2"  
/note="Organ: head; Vector: pOT2; Site 1: EcoRI; Site 2:  
XhoI; Sized fractionated cDNAs were directly ligated into  
pOT2. Plasmid cDNA library."  
ORIGIN  
Alignment Scores:  
Pred. No.: 5,05e-139 Length: 885  
Score: 1273.00 Matches: 227  
Percent Similarity: 91.01% Conservative: 26  
Best Local Similarity: 81.65% Mismatches: 25  
Query Match: 48.22% Indels: 0  
DB: 12 Gaps: 0  
US-09-303-232-6 (1-501) x BG632919 (1-885)  
Qy 54 LeuThrLeuGlnGlnIleAspValAspGluLysAsnGlnLeuLeuIleThrAsnIle 73  
Db 885 CTAACACTCATGCAGATTATCGATGTCACGACGAGAAATCAACTGCTTATAACGAATATT 826  
Qy 74 TrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArgTyrAsnAspSerGluTyrGlyGly 93  
Db 825 TGGCTCAAAATGGAAATGGAACGATATGAATCTTCGATGGAATTCGAGTGAGTTCGGTGT 766  
Qy 94 VallyAspLeuArgIleThrProAsnLysLeuTrpLysProAspValLeuMetTyrAsn 113  
Db 765 GTGGGGATCTGGGNAATTCGGCCATCGCTATCGAAACCGGATGATGATGATACAC 706  
Qy 114 SerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValArgSerGlyGly 133  
Db 705 AGTGGCCGACGAGGGCTTCGATGGAACGATGACGACCAAAATGTGGTTCGCAATTAATGGG 646  
Qy 134 SerCysLeuTyrValProGlyIlePheLysSerThrCysLysMetAspIleAlaTrp 153  
Db 645 AGCTGTCTGTACGTACCGCCAGGTATATTAACTCACTGTAAGATCGACATTCAGTGG 586  
Qy 154 PheProPheAspAspGlnHisCysAspMetLysPheGlySerTyrThrTyrAspGlyAsn 173  
Db 585 TTTCATTCGACGATCAGAGATGTGAATGAAATTTGGTTCGTGACCTACCATGGGTTT 526  
Qy 174 GlnLeuAspLeuValLeuLysAspGluAlaGlyGlyAspLeuSerAspPheIleThrAsn 193  
Db 525 CAGTTGGACCTGCAGTTGCAGACGAGCTGGTGGCAGATTTCTAGCTTTATTAACCAAT 466  
Qy 194 GlyGluTyrTyrLeuIleGlyMetProGlyLysAsnThrIleThrTyrAlaCysCys 213  
Db 465 GCGGAATGGGACTTGTAGTGTGCCGGTAAACGAATGAAATCTACTATAATTTGCTGC 406  
Qy 214 ProGluProTyrValAspValThrPheThrIleMetIleArgArgThrLeuTyrTyr 233  
Db 405 CCAGAACTTATATTGACATAACATTCGCCCATTTTGTATAGGCGCAAAACGTTGTACTAT 346  
Qy 234 PhePheAsnLeuIleValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThr 253  
Db 345 TTTTTCATCTGATTGTGGCGTGGTACTATCGCTCCATGGCATGCTAGGGTTTACA 286  
Qy 254 LeuProProAspSerGlyGluLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThr 273  
Db 285 CTGCCACCATCTTGTGTGAAGAGCTTCGCTTGGAGTTTCAATTTCTATTATCGCTTACA 226  
Qy 274 ValPheLeuAsnLeuValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeu 293  
Db 225 GTCTTCTCAACATGGTGGCGGAAACAATCCGCGGACCTCCGATGGGTACCGCTGCTC 166  
Qy 294 GlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValValLeuThrValVal 313  
Db 165 GGAACCTTATTTCAATTCATATGTTATGTTGGCTCATCAGTTGTGTCAACCATATT 106  
Qy 314 ValLeuAsnTyrHisArgThrAlaAspIleHisGluMetProGlnTrpIle 331  
Db 105 GTCTCAATTTATCATCATAGAAATCCAGATACGATGAAATGAGTGAATGGTA 52  
RESULT 2  
AK053497 1864 bp mRNA linear HTC 20-SEP-2003  
AK053497 Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched  
LOCUS library, clone: E130103E14 product: NEURONAL NICOTINIC ACETYLCHOLINE  
DEFINITION RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.  
ACCESSION AK053497  
VERSION AK053497.1 GI:26343494  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE  
1 Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE  
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE  
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multipillar sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
REFERENCE  
4 The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)



Db	1075	ACACACACACGATGCCACTTGGGTCAAGGCTGTGTTTGTGAACCTTCTCCCGAGGTC	1134
QY	343	LeuArgMetSerArgProGlyLysLysIleThrArgLysThrIleMetMetAsnThrArg	362
Db	1135	ATGTTTATGACTAGGCCAACACGACCGAGGAAGACGCCCAAGACGAGGAACCTTCTAC	1194
QY	363	MetArgGluLeuGlu-----LeuLysGluArgSerLysSerLysLeuLeuAla	378
Db	1195	GGTGCTGAGCTCTCAAACTGAAGTGTCTCAGCGTCGACACTCCAAAGACTGC-----	1248
QY	379	AsnValLeuAspIleAspAspPheArgHisGlyProProProAsnSerThrAla	398
Db	1249	-----AAGGAGGCTTACCCTCCCAAGATGGGACATGT	1281
QY	399	SerThrGlyAsnLeuGlyProGlyCysSerIlePheArg-----ThrAspPheArg	415
Db	1282	GGCTAT-----TGCCACCACCGTAGGTAGGTAATAATCTCAAAATTTTCAGT	1323
QY	416	ArgSerPheValArgProSerThrMetGluAspValGlyGly-----GlyLeuGlySer	433
Db	1324	GCCAACTCTCAAGAGCTCCAGTCTGTGAGTCTGTGATGCTGTGTTGCTCCTCTGCT	1383
QY	434	HisHisArgGluLeuHisLeuIleLeuArgGluLeuGlnPheIleThrAlaArgMetLys	453
Db	1384	CTGTCACCAAGAAATCAAGAGACCTCCAAAGTGTGAAGTATATTGCTGAAATATGAAA	1443
QY	454	LysAlaAspGluGluAlaGluLeuIleSerAspTrpLysPheAlaAlaMetValValAsp	473
Db	1444	GCACAGAATGTAGCAAGAGATTCGAAGATGTTGGAAGTATGTGTCATGGTGAATTGAT	1503
QY	474	ArgPheCysLeuPheValPheThrLeuPheThrIleLeuAlaThrValAlaValLeuLeu	493
Db	1504	CGCATTTTCTCTGGGTTTTTCATCTGTGTGCAATTTAGGAATCGAGGATTATTCTG	1563
RESULT 3			
AK051730			
LOCUS	AK051730	2916 bp	mRNA linear HTC 20-SEP-2003
DEFINITION	Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length		
	enriched library, clone:D130068A06 product:NEURONAL NICOTINIC		
	ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.		
ACCESSION	AK051730		
VERSION	AK051730.1	GI:26342155	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1	Carninci, P. and Hayashizaki, Y.	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,	
AUTHORS	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to		
	prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,	
AUTHORS	Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,		
TITLE	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, I., Harada, A.,		
JOURNAL	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		
MEDLINE	Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,		
PUBMED	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,		
REFERENCE	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
AUTHORS	RIKEN integrated sequence analysis (RISA) system-384-format		
TITLE	sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
11076861			
4			
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the		
AUTHORS	FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 685-690 (2001)		
REFERENCE	5		
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research		
TITLE	Group Phase I & II Team.		
JOURNAL	Analysis of the mouse transcriptome based on functional annotation		
MEDLINE	Nature 420, 563-573 (2002)		
PUBMED	6 (bases 1 to 2916)		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,		
TITLE	Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,		
JOURNAL	Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,		
MEDLINE	Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,		
PUBMED	Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,		
REFERENCE	Koyama, M., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,		
AUTHORS	Okazaki, Y., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,		
TITLE	Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,		
JOURNAL	Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,		
MEDLINE	Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,		
PUBMED	Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,		
REFERENCE	Muramatsu, M. and Hayashizaki, Y.		
AUTHORS	Direct Submission		
TITLE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of		
JOURNAL	Physical and Chemical Research (RIKEN), Laboratory for Genome		
MEDLINE	Exploration Research Group, RIKEN Genomic Sciences Center (GSC),		
PUBMED	RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,		
REFERENCE	Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,		
AUTHORS	URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,		
TITLE	Fax: 81-45-503-9216)		
JOURNAL	cDNA library was prepared and sequenced in Mouse Genome		
MEDLINE	Encyclopedia project of Genome Exploration Research Group in Riken		
PUBMED	Genomic Sciences Center and Genome Science Laboratory in RIKEN.		
REFERENCE	Division of Experimental Animal Research in Riken contributed to		
AUTHORS	prepare mouse tissues.		
TITLE	Please visit our web site for further details.		
JOURNAL	URL: http://genome.gsc.riken.go.jp/		
MEDLINE	URL: http://fantom.gsc.riken.go.jp/.		
PUBMED	Location/Qualifiers		
FEATURES	1. .2916		
source	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="FANTOM.DB:D130068A06"		
	/db_xref="MGI:2420066"		
	/clone="D130068A06"		
	/tissue type="spinal ganglion"		
	/clone_lib="RIKEN full-length enriched mouse cDNA library"		
	/dev_stage="12 days embryo"		
	170. .1669		
	/note="unnamed protein product; NEURONAL NICOTINIC		
	ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT (SPTR)AAL58471,		
	evidence: FASTV, 99.8%ID, 100%length, match=1497)		
	putative"		
	/codon_start=1		
	/protein_id="BAC34740.1"		
	/db_xref="GI:26342156"		
	/translation="MGVVLPFPPLSMMLVLMMLLPVSAASAEHRLFOYLFEDYNEIIL		
	RPVANSHPVIOFEVSMQVLDVNVQINWLNWKOIWNQYKWKPKSDYQGVF		
	MRVPAKWKPDIVLYNNADGDFQVDDKTKALYGTGVTWIPAIKSSCKIDVTYF		
	PFYQNCMTKFGSDYDKADLDVLIGSSMNLKDYSGEWAIKAPQKHEIKVNC		
	BEIQDITYSYIRLPLFYINLIICLLISFLTVLFPVLPDCKEVTLCISVLLS		
	LVFLVTEIRIPSTSLVPLIGVEYLFVTLSIVITVFLNVHVRTPTTHMPT		
	WVKAVFLMLPRVPMFTPTSTEDAPKTRFYGAEINLNCFSRADSKSCKEGYPCQ		
	DGTCYCHHREVKISNFSAITRSSSSVDAVLSALSPEIKAIQSVKVIENMK		
	AQNAKEIQDDQKYVAMVIDRFLVFWFLVLCILGTLGFLQPLMARDDT"		
	2898. .2903		
	/note="putative"		
	2916		
	polya_signal		
	polya_site		



[illegible]

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kura, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system—384-format sequencing pipeline with 384 multipillar sequencer Genome Res. 10 (11), 1757-1771 (2000)

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 3126)









## ORIGIN

## Alignment Scores:

Pred. No.: 6,51e-97 Length: 2296  
 Score: 922.50 Matches: 190  
 Percent Similarity: 56.44% Conservative: 86  
 Best Local Similarity: 38.85% Mismatches: 178  
 Query Match: 34.94% Indels: 35  
 DB: 14 Gaps: 7

US-09-303-232-6 (1-501) x CD013901 (1-2296)

QY 20 GlyProHis-----GluLeuArgLeuLeuAlaLeuLeuAlaAsnTyrAsnThr 36  
 Db |||||  
 Db 332 GGCTCGCATACCGAGACTGAGGACCGGCTCTTCAAACACCTCTTCGGGGGTACACCGC 391  
 QY 37 LeuGluArgProValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThrLeu 56  
 Db |||||  
 Db 392 TGGCGCGCGCGCGGTCGCAACACTTCAGACGTGGTGATTTGGCGCTTTGGACTGTCCATC 451  
 QY 57 GlnGlnIleLeuAspValAspGluLeuAsnGlnLeuLeuThrAsnIleThrLeuSer 76  
 Db |||||  
 Db 452 GCTCAGCTCATCGATGGTGGATGAGAACACAAATGATGACCAACACGCTTGGCTAAAA 511  
 QY 77 LeuGluTrpAsnAspTyrAsnLeuArgTrpAsnAspSerGluTyrGlyGlyValLeuAsp 96  
 Db |||||  
 Db 512 CAGGAGTGGAGCGCACTACAACTGCGTGGAAACCCCGCTGATTTGGCAACATCATCT 571  
 QY 97 LeuArgIleThrProAsnLeuLeuTrpLysProAspValLeuMetTyrAsnSerAlaAsp 116  
 Db |||||  
 Db 572 CTGAGGTCCCTCTGAGATGATCTGGATCCCGACATTTCTTACAAACAATGCGAT 631  
 QY 117 GluGlyPheAspGlyTyrGlnThrAsnValValValArgSerGlyGlySerCysLeu 136  
 Db |||||  
 Db 632 GGGAGTTTTCAGTGCACCCACATGACCAAGGCCACCTCTTCTCCAGGGCACTGTGCAC 691  
 QY 137 TyrValProGlyLeuPheLysSerThrCysLysMetAspIleAlaTrpPheProPhe 156  
 Db |||||  
 Db 692 TGGGTGCCCGCCCATCTACAAGAGCTCTGAGCATCGACGTCACCTCTTCCCTTC 751  
 QY 157 AspAspGlnHisCysAspMetLysPheGlySerTyrTrpAspGlyAsnGlnLeuAsp 176  
 Db |||||  
 Db 752 GACCAAGCAACTGCAAGATGAAGTTGGTCTCGGACTTATGCAAGGCCAAGATCGAC 811  
 QY 177 LeuValLeuLysAspGluAlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGluTrp 196  
 Db |||||  
 Db 812 CTGGAGCATGGAGCAGACT---GTGGACCTTAAGGACTTACGGGAGCGCGCGAGTGG 868  
 QY 197 TyrLeuIleGlyMetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGluPro 216  
 Db |||||  
 Db 869 GCCATCGTCAATGCCCGGCACCTACAAGCAAGAGTACGACTGCTGCGCCGAGATC 928  
 QY 217 TyrValAspValThrPheThrIleMetIleArgArgArgThrLeuTyrTrpPheAsn 236  
 Db |||||  
 Db 929 TACCCGCGAGTCACCTACCGCTTCGTCATCCGCGCGCTCGCGCTCTTACACCATCAAC 988  
 QY 237 LeuIleValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProPro 256  
 Db |||||  
 Db 989 CTATCATCTCCCTGCTGCTCATCTCTCCCTACGTGCTGCTTCTTACCTGCCCTCC 1048  
 QY 257 AspSerGlyGluLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeu 276  
 Db |||||  
 Db 1049 GACTGCGGCGAAGATCATCGCTGTGTCATTTGGTGTCTGCTCATCACTACCGTCTTCCTG 1108  
 QY 277 AsnLeuValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyr 296  
 Db |||||  
 Db 1109 CTGCTCATCAGATCATATCCGTCACCTCGCTGCTCATCCGCTCATCGCGCTATCGCGAGTAC 1168  
 QY 297 PheAsnCysIleMetPheMetValAlaSerValValLeuThrValValValLeuAsn 316  
 Db |||||  
 Db 1169 CTGCTGTTCCACCATGATCTTCGTACCCCTGTCTCATCTGTCATCAGCTCTTCGTGCTCAAT 1228  
 QY 317 TyrHisHisArgThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPheLeu 336

Db 1229 GTGCACCACCGCTCCCCAGCACCCACACCATGCCCCACCTGGGTGCGGGGCGCCCTCTG 1288  
 QY 337 GlnTrpLeuProTrpIleLeuArgMetSerArgProGlyLysLysIleThrArgLysThr 356  
 Db |||||  
 Db 1289 GGCTGTGTCGCCCGGTGCTTCTCATGAACCGGCGCCACCACCGGTG----- 1336  
 QY 357 IleMetMetAsnThrArgMetArgGluLeuGluLeuLysGluArgSerSerLysSerLeu 376  
 Db |||||  
 Db 1337 -----GAGCTCTGCCACCCCTCAGCTGAAGCTCAGCCCTCTTATCACTGG 1384  
 QY 377 LeuAlaAsnValLeuAsp-----IleAspAspPheArg--- 388  
 Db |||||  
 Db 1385 CTGAGAGCAACGTGGATCGCGAGAGAGGAGTGGTGGAGGAGGAGACAGATGG 1444  
 QY 389 -----HisGlyProProProAsnSerThrAlaSerThrGlyAsnLeuGly 404  
 Db |||||  
 Db 1445 GCATGTGCAGTCTATGTGGCCCTCTGTGGGACCCCTCTGCAGCCACGCCCTGCAC 1504  
 QY 405 ProGlyCysSerIlePheArgThrAspPheArgSerPheValArgProSerThrMet 424  
 Db |||||  
 Db 1505 TCTGGGGCC-----TCAGGTCCCAAGGCTGAGGCTCTG 1537  
 QY 425 GluAspValGlyGlyGlyLeuGlySerHisHisArgGluLeuHisLeuIleLeuArgGlu 444  
 Db |||||  
 Db 1538 CTGACGAGGGTGAGCTGCTATCACCCAC-----ATGCAGAGGCACCTGAAGGT 1591  
 QY 445 LeuGlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuIleSerAsp 464  
 Db |||||  
 Db 1592 GTGCACCTACATGCCGACCCACCTGCGGTCTGAGGATGCTGACTCTCGGTGAGGAGGAC 1651  
 QY 465 TrpLysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThr 484  
 Db |||||  
 Db 1652 TGAAGATGATGTCCTGCTATCGACAGGATCTCTCTGCGCTGTTATCATCGTCTGC 1711  
 QY 485 IleIleAlaThrValAlaValLeuLeu 493  
 Db |||||  
 Db 1712 TTCCTGGGACCATCGCCTCTTCTG 1738

## RESULT 9

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## PUBMED

## AUTHORS

## COMMENT

## FEATURES

## SOURCE

AY402878 1454 bp DNA linear GSS 15-DEC-2003  
 Mus musculus CHRN4 gene, VIRTUAL TRANSCRIPT, partial sequence,  
 genomic survey sequence.  
 AY402878  
 AY402878.1 GI:39758861  
 GSS.  
 Mus musculus (house mouse)  
 Mus musculus  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1454)  
 Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 Science 302 (5652), 1960-1963 (2003)  
 14671302  
 2 (bases 1 to 1454)  
 Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 Location/Qualifiers  
 1..1454  
 /organism="Mus musculus"



Qy	23	GlulysArgLeuLeuAsnAlaLeuLeu-----AlaAsnTyrAsnThrLeuGluArgPro	40
Db	24	GAGGAGAAGCTGATGGATGATCTCTCTGAACAAAACTCGCTACAAACAACTGATCGCCCA	83
Qy	41	ValAlaAsnGluSerGluProLeuGluValAlaArgPheGlyLeuThrLeuGlnIleIle	60
Db	84	GCCACCAGCTCTCCAGCTCATCTCCATCGCGCTGGAGCTATCACTGTCCAGCTCATC	143
Qy	61	AspValAspGluLysAsnGlnLeuLeulleThrAsnIleTrpLeuSerLeuGluTrpAsn	80
Db	144	AGTGTGAATGAGCGAGAACAGATCATGATGACCCACAGCATCTGGCTAAAAACAGGAATGAGCT	203
Qy	81	AspTyrAsnLeuArgTrpTrpAsnAspSerGluTyrGlyValLysAspLeuArgIleThr	100
Db	204	GACTACCGCCCTGGCGTGAACAGCTCCCTGCTATGAAGGGGTGAACATTCAGGAGATTCCT	263
Qy	101	ProAsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp	120
Db	264	GCAAGCGAGCTGTGTTCCCTTGACATTTGTTGTGTACAAATGCCGATCGGACCTATGAG	323
Qy	121	GlyThrTyrGlnThrAsnValValAlaArgSerGlyGlySerCysLeuTyrValProPro	140
Db	324	GTGTCTGTCTACACCAAGCTGATCTGTGGTTCCTCAACGGCAGCATCCAGTGGTGGCCCT	383
Qy	141	GlyIlePheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHis	160
Db	384	GCCATCTCAAGAGTGCCTGCAAGATAGAGGTGAAGACTTTCCCTTCGACCAGCAGAAC	443
Qy	161	CysAspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLys	180
Db	444	TGCACCTTCAATTTTCGTTCTCGTCTTGACCTATGACCACACAGAGATGACATGTGTTCTAA	503
Qy	181	AspGluAlaGlyLysLeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGly	200
Db	504	TCACCCACA--GCATCATGATGATGACTTATACCCCGCGGTGAATGGACATTTGTGGCC	560
Qy	201	MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGlu-----ProTyrVal	218
Db	561	CTCCCGGTTCGAAG--ACAGTGAC-----CCACGAGACCCACGATATGTG	605
Qy	219	AspValThrPheIleMetIleArgArgArgThrLeuTyrTyrPheAsnLeuIle	238
Db	606	GACGTGACCTATGACTTTCATCATCAAGCGCAAGCCACTCTTCTACACCATCAATCTGATC	665
Qy	239	ValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSer	258
Db	666	ATCCCTTGTGTCTATCATCGTGTGGTATCTCTGTCTTCTACCTGCCCTCCGACTGT	725
Qy	259	GlyGluLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeu	278
Db	726	GGGGAGAAAGATGACGCTCTGCATCTCAGTGCTCTGTCGACTCACATTCCTCTCTGCTGCTC	785
Qy	279	ValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsn	298
Db	786	ATCTCCAAGATCTGCTGCCCTCCACCTCCCTTGACATTCCTCATTCGCAAGTACCTCTTG	845





```
Db      1046 ATCTCAAGATTGTGCTCCACCTCCCTCGAGCTACCGTGTGGGAAAGTACCTCATG 1105
Qy      299 CysIleMetPheMetValAlaSerSerValValLeuThrValValValLeuAsnTyrHis 318
Db      1106 TTCACCATGGTGCTAGTACCTTCCTCCATCGTCACTAGCGTGTGTGTCTCAATGTGCAC 1165
Qy      319 HisArgThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPheLeuGlnTrp 338
Db      1166 CACCGTGTGCTACCGACACACATGGCCGCTTGGTCAAGTGTCTTCTTGGAGAAG 1225
Qy      339 LeuProTrpIleLeuArgMetSerArgProGlyLysLysIleThrArgLysThrIleMet 358
Db      1226 CTGCCACCTCTCTCTCTGCGACGACCGCTGTGACGCTCAACGCTGCCTGCGC 1285
Qy      359 MetAsnThrArgMetArgGluLeuGluLysGluArgSerSerLysSerLeuLeuAla 378
Db      1286 TTGCGAAGCGCGCGGAGCACTGAG-----GGGGCAGGCACACATTATC--- 1330
Qy      379 AsnValLeuAspIleAspAspPheArgHisGlyProProProProAsnSerThrAla 398
Db      1331 -----TTCGCGAAGGCTCTGCGAGCTGACCGCTGACCATGTACTGTC 1366
Qy      399 SerThrGlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgSerPhe 418
Db      1367 -----TTT 1369
Qy      419 ValArgProSerThrMetGluAspValGlyGly----- 429
Db      1370 GTCAATCTCTGCATCAATGCGAGGCTTGTGCTGGGCTTTCCAGGCTGAGCCGCTGCAGCC 1429
Qy      430 GlyLeuGlySerHisHisArgGluLeuHisLeuLeuLeuLeuLeuLeuLeuLeuLeu 445
Db      1430 GGCTGGGGCGCTCGATGGGCCCATGCGAGCTGCTGGCTCCGGGAGCTGTGCGAGGTGA 1489
Qy      446 GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuLeuLeuSerAspTrp 465
Db      1490 CGCTTCATTCGGGACCATATGCGAAGTGAAGATGATGACCAAGATGTGAGGAGGATTGG 1549
Qy      466 LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle 485
Db      1550 AAATACCTGTGCATGGTGTGATCGACCGCTGTCTCTGGATCTTTGTCTGTGTCTGTC 1609
Qy      486 IleAlaThrValAlaValLeuLeu 493
Db      1610 TTTCGGACCATGCGATGTCCTG 1633

RESULT 11
LOCUS   AK051742
DEFINITION Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
           enriched library, clone: D130070121 product: cholinergic receptor,
           nicotinic, beta polypeptide 2 (neuronal), full insert sequence.
ACCESSION AK051742
VERSION   AK051742.1 GI:26342173
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS   Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
           Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED    10349636
REFERENCE
AUTHORS   Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
           Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
           prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
PUBMED    11042159
```

```
3      Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
      Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
      Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
      Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
      Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Chata, E., Matsuoka, M.,
      Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
      Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
      RIKEN integrated sequence analysis (RISA) system--384-format
      sequencing pipeline with 384 multicapillary sequencer
      Genome Res. 10 (11), 1757-1771 (2000)
      20530913
      11076861
4      The RIKEN Genome Exploration Research Group Phase II Team and the
      FANTOM Consortium
      Functional annotation of a full-length mouse cDNA collection
      Nature 409, 685-690 (2001)
5      The FANTOM Consortium and the RIKEN Genome Exploration Research
      Group Phase I & II Team.
      Analysis of the mouse transcriptome based on functional annotation
      of 60,770 full-length cDNAs
      Nature 420, 563-573 (2002)
      6 (bases 1 to 4046)
      Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
      Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
      Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
      Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
      Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
      Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Ohno, M., Ohsato, N.,
      Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Sakai, K., Sakazume, N.,
      Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
      Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
      Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
      Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
      Muramatsu, M. and Hayashizaki, Y.
      Direct Submission
      Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
      Physical and Chemical Research (RIKEN), Laboratory for Genome
      Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
      RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
      Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
      URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
      Fax: 81-45-503-9216)
      cDNA library was prepared and sequenced in Mouse Genome
      Encyclopedia Project of Genome Exploration Research Group in Riken
      Genomic Sciences Center and Genome Science Laboratory in RIKEN.
      Division of Experimental Animal Research in Riken contributed to
      prepare mouse tissues.
      Please visit our web site for further details.
      URL: http://genome.gsc.riken.go.jp/
      URL: http://fantom.gsc.riken.go.jp/.
      Location/Qualifiers
          1. 4046
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="C57BL/6J"
             /db_xref="FANTOM,DB:D130070121"
             /db_xref="MGI:2420100"
             /db_xref="taxon:10090"
             /clone="D130070121"
             /tissue type="spinal ganglion"
             /clone_lib="RIKEN full-length enriched mouse cDNA library"
             /dev_stage="12 days embryo"
          209..1714
             /note="unnamed protein product; cholinergic receptor,
             nicotinic, beta polypeptide 2 (neuronal) (MGD|MG1:87891,
             GIN|NM_009602, evidence: BLASTN, 99%, match=1498)
             putative"
             /codon_start=1
             /protein_id="BAC34749.1"
             /db_xref="GI:26342174"

FEATURES
          source
          CDS
```

translation="WARCNSNALLPSGLLWLCGVLGDTDEERLVEHLDPFRYNK  
LIPATNGSELTVQLWVSLAQIISVHREKQITNMVLQWBEYIATWKQEDFDNF  
KXVLPKSHMLPDLVXNADMGVEFYSNVAYDGSIFWLPALYKACKIRVK  
HFPPQDQNTFKRSPYTDIEILVADKSLADDFTPSGWDLIALPERGNEPDD  
STVDITDQILFRKRLPYTNILPCVLITSLAIFLVPFSDCGEMCTLSVIAL  
TVVLLKSLPFLFQDQPHRCARQLRRQRREAGTFLVHSGVNLHRRPPTWMAFP  
KVVFLBEKLPFLQDQPHRCARQLRRQRREAGTFLVHSGVNLHRRPPTQFVN  
PASMGHLAGAFQAFPAALGQDGMGSCGLREAVDGVRFIAHMRSDSDQVRED  
KYAMVLDRLFLFVAVLGVFGTIGMFLQYFQNTYATLFLHSDKAPSSK"

Alignment Scores:			
Pred. No.:	4,31e-95	Length:	4046
Score:	910.50	Matches:	185
Percent Similarity:	57.38%	Conservative:	95
Best Local Similarity:	37.91%	Mismatches:	153
Query Match:	34.49%	Indels:	55
DB:	11	Gaps:	9
US-09-303-232-6 (1-501)	x AK051742	{1-4046}	

QY	23	GluLysArgLeuLeuAsnAlaLeuLeu-----AlaAsnTyrAsnThrLeuGluArgPro	40
Db	293	GAGAGCGGCTGGTGGAGCATCTCTTGGATCTCTCCGCTATAACAAGCTGATCCGTCCA	352
QY	41	ValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnGlnLeuLeile	60
Db	353	GCTACTAATGGCTCTGAGCTGGTGTGACTGTACAGCTCATGGTATCATTTGGCACAGCTCATC	412
QY	61	AspValAspGluLysAsnGlnLeuLeuLeileThrAsnIleTrpLeuSerLeuGluTrpAsn	80
Db	413	AGTGTGCACGAGCGGAGCAGATCATGACCCACCAACGCTGGCTGACCCAGAGTGGGAA	472
QY	81	AspTyrAsnLeuArgTrpAsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThr	100
Db	473	GATTATCGCCTCACATGGAAGCCTGAGGATTTGCACAATATGAAGAAAGTCCGACTCCCT	532
QY	101	ProAsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp	120
Db	533	TCTAAGCACATCTGGCTCCAGATGTGTTCTATACACAATGCTGACGGCATGTACGAA	592
QY	121	GlyThrTyrGlnThrAsnValValValArgSerGlyGlySerCysLeuThrValProPro	140
Db	593	GTCTCTCTTATTCCAAATGCTGTGCTCCATGATGGCAGCATCTTTGGCTACCGCT	652
QY	141	GlyIlePheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHis	160
Db	653	GCCATCTACAGAGCGCATGCAAGATGAGGTGAAGCACTTCGCATTTGACCACGACAA	712
QY	161	CysAspMetLysPheGlySerTyrThrTyrAspGlyAsnGlnLeuAspLeuValLeuLys	180
Db	713	TGCACCATGAAGTTCGGCTCTCTGACCTACGACCGACTGAGATTGACCTGGTGGCTCAA	772
QY	181	AspGluAlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGly	200
Db	773	ACGGATGTG---GCCAGCTGGAGCACTTACACCTAGTGGGAGTGGGACATCATCGCA	829
QY	201	MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGlu-----ProTyrVal	218
Db	830	CTGCCAGCGCCAGCAACGAG-----AACCCAGCACTCCACCTACGCT	874
QY	219	AspValThrPheThrIleMetIleArgArgThrLeuTyrTyrPhePheAsnLeuLeu	238
Db	875	GACATCACTACGACTTCATCATCTGTCGCAACCGCTCTTCTACACCATCAACTCTCATC	934
QY	239	ValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProAspSer	258
Db	935	ATCCCTCTCGTACTCATCACTCGCTGCCCATCTGTCTTCTTACCTGCGCTCAGACTGT	994
QY	259	GlyGluLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeu	278
Db	995	GGTGAAGAAGATGACACTTTGATTTCTGTCTCTCTGGCGTACCGGTGTTCTCGTGTGTC	1054

QY	279	ValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsn	298
Db	1055	ATCTCCAAGATTGTGCCTCCACACCTCCCTCGACCTACCGTGGTGGGAAGTACCTCATG	1114
QY	299	CysIleMetPheMetValAlaSerSerValValLeuThrValValValLeuAsnTyrHis	318
Db	1115	TTCACCAATGGTGTAGTACACTTCTCATGTCTACTAGCGTGTGTGTCTCAATGTGCAC	1174
QY	319	HisArgThrAlaAspIleHisGluMetProGlnThrPileLysSerValPheLeuGlnTrp	338
Db	1175	CACGTTCCGCTACACGACACCATGGCGCTTGGGTCAAGTGGTCTTCTCTGGGAAG	1234
QY	339	LeuProTrpIleLeuArgMetSerArgProGlyLysIleThrArgLysThrIleMet	358
Db	1235	CTGCCCACCTCCTCTTCTGCAGACGCCACCGCTGTGCAGTCAACCCCTGGCG	1294
QY	359	MetAsnThrArgMetArgGluLeuGluLeuLysGluArgSerSerLysSerLeuLeuAla	378
Db	1295	TTGCGAAGCGCGACGCGGAACGTGAG-----GGGCAGCAGCACATATTC---	1339
QY	379	AsnValLeuAspIleAspAspPheArgHisGlyProProProAsnSerThrAla	398
Db	1340	-----TTCCGCGAAGGTCTCTGCAGTGCACCATGTACCTGC	1375
QY	399	SerThrGlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgSerPhe	418
Db	1376	-----TTT	1378
QY	419	ValArgProSerThrMetGluAspValGlyGly-----	429
Db	1379	GTCAATCTGCATCAATGACGGCTTGGCTGGGGCTTCCAGCGTGAAGCCCTGCAGCC	1438
QY	430	GlyLeuGlySerHisArgGluLeuHisLeuLeuArgGlu-----Leu	445
Db	1439	GGTCTGGGGCGCTCGATGGGCCCATCGAGCTGTGGCTCCGGGAAGCTGTGGACGGTGA	1498
QY	446	GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuSerAspTrp	465
Db	1499	CGCTTATTTCGGACCATATGCAAGTGAAGATGATGACAGAGTGTGAGGGAGGATTGG	1558
QY	466	LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle	485
Db	1559	AAATACGTTGCCATGTGATCGACCGCTGTCTCTGGAACCTTGTCTTGTCTGTGTGTC	1618
QY	486	IleAlaThrValAlaValLeuLeu	493
Db	1619	TTTGGGACCAATTGGCATGTTCCTG	1642

## RESULT 12

AL530299	AL530299	1201 bp	mRNA	linear	EST 23-MAY-2003
DEFINITION	AL530299 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens cDNA clone CS0DD007YP05 5-PRIME, mRNA sequence.				
ACCESSION	AL530299				
VERSION	AL530299.2	GI:31068132			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 1201)				
TITLE	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
JOURNAL	Full-length cDNA libraries and normalization				
COMMENT	Unpublished (2001)				
	On Feb 13, 2001 this sequence version replaced an previous				

Contact: genomecope  
 Genomecope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: [segrete@genomecope.cns.fr](mailto:segrete@genomecope.cns.fr), Web : [www.genomecope.cns.fr](http://www.genomecope.cns.fr)  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 7646.r For more information about this cluster, see  
<http://www.genomecope.cns.fr/>

cgi-bin/cluster.cgi?seq=CS0DD007CH03QPI&cluster=7646.i.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0DD007CH03QPI.

# FEATURES

Location/Qualifiers  
1..1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DD007YP05"  
/tissue="NEUROBLASTOMA COT 50-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"  
/note="First strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with NotI and cloned into the NotI and EcoRV  
sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Alignment Scores:  
Pred. No.: 8.9e-96 Length: 1201  
Score: 909.00 Matches: 181  
Percent Similarity: 73.19% Conservatives: 51  
Best Local Similarity: 57.10% Mismatches: 79  
Query Match: 34.43% Indels: 8  
DB: 9 Gaps: 2

US-09-303-232-6 (1-501) x AL530299 (1-1201)

QY 8 LeuAlaLeuLeuAla---LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26  
DB 152 CTGGCGCTGGCGGCTGCTCTCTGACGTGCTCTGCAAGCGAG-TWCCAGAGGAAGCTT 210  
QY 27 LeuAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSerGlu 46  
DB 211 TACAAGGAGCTGGTCAAGAACTACAAATCCCTTGGAGAGCGCGCTGCAATGATCGGAA 270  
QY 47 ProLeuGluValArgPheGlyLeuThrLeuGlnGlnLeuAspValAspGluLysAsn 66  
DB 271 CCACCTACCGTCTACTTCTCTGAGCTCTGACATGATGATGATGATGATGATGATGATGAT 330  
QY 67 GlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 86  
DB 331 CAAGTTTAAACCAACCAATTTGGTCTGCAAGTCTTGGACAGATCACTATTATACAGTGG 390  
QY 87 AsnAspSerGluTyrGlyGlyValLysAspLeuArgGlnLeuThrProAsnLysLeuTrpLys 106  
DB 391 AATGTGTGAGATATCCAGGGGTGAAGCTGTTTCCAGATGCGGCGAGATTTGGAAA 450  
QY 107 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyTyrTyrGlnThrAsn 126  
DB 451 CCAGACATTTCTCTATACAGTGTGATGAGCGCTTTCGACCACTTCCACACTAAC 510  
QY 127 ValValValArgSerGlyGlySerCysLeuTyrValProProGlyLeuPheLysSerThr 146  
DB 511 GTGTGTGTGAATTTCTTCTGCGCATTTGCCAGTWCCTTCCAGATGCGGCGAGATTTGCC 570  
QY 147 CysLysMetAspLeuAlaTrpPheProPheAspGlnHisCysAspMetLysPheGly 166  
DB 571 TGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630  
QY 167 SerTyrThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyGlyAsp 186  
DB 631 TCCTGTCTTACGAGGCTGGTCTTGGATCTGACATGACAGAG-----GCAGAT 681  
QY 187 LeuSerAspPheLeuThrAsnGlyGluTyrTyrLeuLeuGlyMetProGlyLysLysAsn 206  
DB 682 ATCAGTGGCTATATCCCAATGGAATGGACCTAGTGGAAATCCCGGCAAGAGAGT 741  
QY 207 ThrLeuThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMetIle 226  
DB 742 GAAAGGTTTATGAGTCTGCAAGAGCCCTACCCGATGTCACCTTTCAGATGACCATG 801  
QY 227 ArgArgArgThrLeuTyrTyrPhePheAsnLeuLeuValProCysValLeuLeuSerSer 246

DB 802 CGCGCAGGAGCGCTTACTATGCGCTCAACCTGCTGATCCCTGTGTGCTCACTCGCC 861  
QY 247 MetAlaLeuLeuGlyPheThrLeuProAspSer-GlyGluLysLeuThrLeu-GlyV 266  
DB 862 CTGCGCTGCTGTGTCTCTGCTTCTGCGAGATTCGGGGGAGAGATTTCCTTKGGGA 921  
QY 266 alThrLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnV 286  
DB 922 TAACAGTGTACTCTCTTACCGTCTTACGCTGCTGCTGCTGAGATCATGCCCGCAA 981  
QY 286 alSerAspAlaLeuProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAla 306  
DB 982 CATCCGATTCGGTACCATTCGAGCCCGAGTCTTCCGACGACCATCATCTCGGGCT 1041  
QY 306 erSerValValLeuThrValValValLeuAsnTyrHisHisArgThr 321  
DB 1042 TTGG-GTGGTTGKRMGGTGTATCGTGTGCAGTACCAACACCGACGACC 1087

RESULT 13  
AY406230 1374 bp DNA linear GSS 15-DEC-2003  
LOCUS  
DEFINITION Homo sapiens CHRNA1 gene, VIRTUAL TRANSCRIPT, partial sequence,  
Genomic survey sequence.  
ACCESSION AY406230  
VERSION AY406230.1 GI:39762204  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
1 (bases 1 to 1374)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.

TITLE  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL  
Science 302 (5652), 1960-1963 (2003)  
PUBMED  
14671302

REFERENCE  
AUTHORS  
2 (bases 1 to 1374)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.

TITLE  
Direct Submission  
JOURNAL  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA

COMMENT  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source  
1..1374  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>1374  
/gene="CHRNA1"  
/locus\_tag="HMC2488"

gene  
ORIGIN

Alignment Scores:  
Pred. No.: 3.3e-95 Length: 1374  
Score: 905.00 Matches: 182  
Percent Similarity: 56.02% Conservatives: 97  
Best Local Similarity: 36.55% Mismatches: 149  
Query Match: 34.28% Indels: 70  
DB: 29 Gaps: 8

US-09-303-232-6 (1-501) x AY406230 (1-1374)  
QY 4 MetLeuAlaLeuAlaLeuLeuAlaLeuLeuProValSerGluGlnGlyProHisGlu 23  
DB 22 CTGCTTTAGCTTTGCTCAGCTGGCTCGTCTCGGCTCCGAA-----CATGAG 72



of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, A., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku, A., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Havashizaki, Y.

TITLE

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

COMMENT

cdNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

prepare mouse tissues.  
Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>  
URL: <http://fantom.gsc.riken.go.jp/>

FEATURES  
SOUND

```

1. .2340
location=Ga.affinis
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9330165I16"
/db_xref="MGI:2398424"
/db_xref="taxon:10090"
/clone="9330165I16"
/sex="male"
/tissue_type="diencephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
116. .2005
/note="unnamed protein product; cholinergic receptor,
nicotinic, alpha polypeptide 4 (MGD)MGI:87888,
GB|NM_015730, evidence: BLASTN, 99%, match=1946)
out_type="

```

## ORIGIN

Alignment Scores:

Pred. No.:	5,926-93	Length:	2940
Score:	890.50	Matches:	207
Percent Similarity:	48.05%	Conservative:	88
Best Local Similarity:	33.71%	Mismatches:	195
Query Match:	33.73%	Indels:	125
NA:	11	Gaps:	11

US-09-303-232-6 (1-501) x AK034228 (1-2940)

QY	3	ProMetLeuAlaAlaLeuAlaLeuLeu-----AlaLeuLeuProValSer 17
DB	143	CGCGCGTCTGCTGCTCGCGCTCTTAGGACCGGCTCTTGCCTGTAGC 202
QY	18	-----GluGlnGlyProHis---GluLysArgLeuLeuAsnAlaLeuAlaAsn 33
DB	203	AGGCACATGAGACCCGGGCCCATCGCGAGGAGCGCTCTCTGAAGAGACTCTTCTCTGC 262
QY	34	TyrAsnThrLeuGluArgProValAlaAsnGluSerGluProLeuGluValArgPheGly 53
DB	263	TACACAAAGTGTCTCGGCGATGACCAATATCTCAGATGTGGTCTGTTCGCGCTTGGC 322
QY	54	LeuThrLeuGlnGlnIleLeuAspValAspGluLysAsnGlnLeuLeuLeuThrAsnIle 73
DB	323	TTTCGATTGCTCAGCTCATCTGATGTGGATGAGAAAAACAGATGATGACGACCAACGTG 382
QY	74	TrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArgTrpAsnAspSerGluTyrGlyGly 93
DB	383	TGGGTGAAGCAGAGTGGATGACTACAACTCGCGCTGGGACCCCTGGTGACATCGAGAT 442
QY	94	ValLysAspLeuArgIleThrProAsnLysLeuTrpLysProAspValLeuMetTyrAsn 113
DB	443	GTCACTCCATCGCATCCATCTGAACCTCATCTGGAGGCTGCACATGCTCTCTACAAC 502
QY	114	SerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValValArgSerGlyGly 133
DB	503	AACGCGACGGGACTTTGCAGTCACCCCACTAACCAAGCCCACTGTTCTTATGATGGG 562
QY	134	SerCysLeuTyrValProProGlyIlePheLysSerThrCysLysMetAspIleAlaTrp 153
DB	563	CGTGTGCAGTGGACACCCCGGCCATCTATAAGAGCTCTCTGAGCATCGACGTCACTTC 622
QY	154	PheProPheAspAspGlnHisCysAspMetLysPheGlySerTrpThrTyrAspGlyAsn 173
DB	623	TTCCCTCTCCACGACGAACTGTACCATGAAGTTTGGTCTCTGGACCTACGCAAGGCC 682
QY	174	GlnLeuAspLeuValLeuLysAspGluAlaGlyLysAspLeuSerAspPheIleThrAsn 193
DB	683	AAGATTGACTTGTGTG---AGCATGCAACCGGTGTGGACCACTGACCTCTGGGAAAGT 739
QY	194	GlyGluTrpTyrLeuIleGlyMetProGlyLysLysAsnThrIleThrTyrAlaCysCys 213
DB	740	GGGAGTGGTCAATTGTGGATGCGGTGGGCACTACACACGAGGAGTATGATGCTGT 799
QY	214	ProGluProTyrValAspValThrPheThrIleMetIleArgArgArgThrLeuTyrTyr 233
DB	800	GCCGAGATCTATCTGCATCACCTACGCCCTTCATCATCGCGCACTGCGACCTGTCTAC 859
QY	234	PhePheAsnLeuIleValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThr 253
DB	860	ACCATCAACCTTATCATCCGCTGCTCTCATCTCTGCTCCCGTGGTGTCTTCTAT 919
QY	254	LeuProProAspSerGlyGluLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThr 273
DB	920	CTGCCCTCGGAGTCGGCGAAGGTACACGCTGTGCATCTCGGTGTGCTTTCTCTCACC 979
QY	274	ValPheLeuAsnLeuValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeu 293
DB	980	GTCTTCTGCTGCTCATCACCGAGATCATCCCGTCCACCTCGGTGTGTCTCCCGCTCATC 1039
QY	294	GlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValValLeuThrValVal 313
DB	1040	GGCGAGTACCTGTCTTTCACCATGATCTTCGTACCCCTCTCCATTTGCATCAGCTTTC 1099
QY	314	ValLeuAsnTyrHisHisArgThrAlaAspIleHisGluMetProGlnTrpIleLysSer 333
DB	1100	GTGCTCATGTACACCCCGCTCACCACGACACACACCATGCGCGCTGGGTGGCGAGA 1159
QY	334	ValPheLeuGlnTrpLeuProTrpIleLeuArgMetSerArgProGly-----LysLys 351

Db	1160	GTCTTCTCGACATGTGCGCCGCTCTCTTTCATGAAGCGCCCATCTGGTCAAGAC	1219
Qy	352	IleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeuGluLeu	368
Db	1220	AACGCGGAGACTTATCGAATCCATGACACAAAGATGCGCAAGCCCTCGTTCTGGCCA	1279
Qy	369	LysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAsp	382
Db	1280	GAGCCTGAGAGTGAGCCCGGCATCTTGGGTGACATCTGCAACCAAGACCTGTCACTGCC	1339
Qy	383	-----IleAsp-AspAspPheArgHisGlyProPr	392
Db	1340	CCAACTTTCGAACCCGATGGACACAGCAGTTCAGACCCAGCCTCATCGAGTCAACC	1399
Qy	392	OpToProAsnSer	396
Db	1400	TCCACAAAGGTCCCTGACTTGAAGACATCAGAGTTTGAGAAGGCCAGTCCCTGTCCATCA	1459
Qy	396	-----	396
Db	1460	CCTGGCTTTGTCAACCACCAATAGCAGTGGGGCCCGAGTGTCTATCAAAAGCCAGGTCC	1519
Qy	397	-----ThrAlaSerThrGlyAs	402
Db	1520	CTGAGCGTCAGCATGTGCCAGTCCCGAAGACGACCGAGGCGACATCCGCTGCGCG	1579
Qy	402	nLeuGlyProGlyCysSerIlePheArgThrAspPheArgArgSerPheVal	419
Db	1580	TCTCGAGTATCCAGTACTGTGTTTCC-CAAGATGGAGCTGCTTCCCTGACTGAGACAA	1638
Qy	419	-----	419
Db	1639	GCCCACTGGTCCCGACGCTCCCTGGAAGACCCGTCATCCAGCTTCCAGTGTCAAGCA	1698
Qy	420	-----ArgProSerThrMetGluAspVal	427
Db	1699	GACCTCTCCATGCAATGCATGCACGAGGAACCATCTCCTGTGTGCCCCATCACTGTGCT	1758
Qy	428	-----GlyGlyGlyLeuGlySerHisArgGluLeuHisLeu	440
Db	1759	CAAGGCTGGAGGCCAACCAAGACCTCCCAACACCTGCCCTGTCAACGCCCTGACAG	1818
Qy	441	IleLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAlaAlaGluAlaGlu	460
Db	1819	GGCAGTAGAAGGCGTCCAGTACATTTGCAGACCACTCAAGGCGAAGACACAGACTTCTC	1878
Qy	460	uLeuIleSerAspTrpLysPheAlaAlaMetValValAspArgPheCysLeuPheValPh	480
Db	1879	GGTGAAGGAGGACTGGAAATCATGGGCAATGTCATTGACCGAATCTTCTCTGGATGTT	1938
Qy	480	eThrLeuPheThrIleIleAlaThrValAlaValLeuLeu	493
Db	1939	CATCATGTTCTGCTTCTGGGCATCTGGGACTTCTTCTG	1978
RESULT 15			
AK083157			
LOCUS	AK083157	3230 bp	mRNA
DEFINITION	Mus musculus adult male hippocampus cDNA, RIKEN full-length enriched library, clone: C63019M18 product: cholinergic receptor, nicotinic, alpha polypeptide 4, full insert sequence.		
ACCESSION	AK083157		
VERSION	AK083157.1	GI:26350296	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		







```
Db 1739 GGCAGTAGAAGGCGTCCAGTACATTGCAGACCCCTCAAGGCGAAGACACAGACTTCTC 1798
Qy 460 uLeuIleSerAspTrrPheAlaMetValValAspArgPheCysLeuPheValPh 480
Db 1799 GGTGAAGGAGGACTGGAAATACGTGCCATGGTCAATTGACAGAATCTTCTCTGGATGTT 1858
Qy 480 eThrLeuPheThrIleAlaThrValAlaValLeuLeu 493
Db 1859 CATCATTTGTGCTTCTGGGCACTGTGGGACTCTTCCTG 1898
```

Search completed: May 9, 2004, 04:21:03  
Job time : 3532.21 secs

***This Page Blank (uspto)***